

SISTEM REKOMENDASI DESTINASI WISATA BERBASIS *CONTENT-BASED FILTERING* DAN ANALISIS FITUR GEOSPASIAL

Muhammad Arya Widika¹, Putri Salsabilla Susilo², Andhika Ibnu Ramadhan³, Syifa Nur Rakhmah⁴, Findi Ayu Sariasih⁵, Imam Sutoyo⁶

^{1,2,3,4,5,6}Universitas Bina Sarana Informatika
Jalan Kramat Raya No.98, Senen, Jakarta Pusat

Email : ¹aryawidika@gmail.com, ²putrisalsabillasusilo@gmail.com, ³ibnuramadhan2200@gmail.com, ⁴syifa.snk@bsi.ac.id, ⁵findi.fav@bsi.ac.id, ⁶imam.ity@bsi.ac.id

^{1,2,3,5}Fakultas Teknik dan Informatika, Program Studi Informatika, Universitas Bina Sarana Informatika, Jakarta, Indonesia

^{4,6}Fakultas Teknik dan Informatika, Program Studi Teknologi Informasi, Universitas Bina Sarana Informatika, Jakarta, Indonesia

ABSTRAK

Penelitian ini mengembangkan sistem rekomendasi destinasi wisata berbasis *Content-Based Filtering* yang diintegrasikan dengan analisis fitur geospasial untuk meningkatkan relevansi dan konteks rekomendasi bagi pengguna. Sistem ini dirancang untuk menjawab kelemahan pada sebagian besar sistem rekomendasi wisata yang masih berfokus pada aspek popularitas dan penilaian umum tanpa mempertimbangkan kedekatan lokasi maupun preferensi personal. Dataset yang dimanfaatkan berasal dari Tourism in Indonesia (Kaggle) dengan area penelitian difokuskan pada wilayah Jakarta dan Bandung. Dalam proses pengembangan, fitur teks diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), sedangkan tingkat kemiripan antar destinasi dihitung melalui *Cosine Similarity*. Selain itu, jarak geografis antar lokasi dianalisis menggunakan rumus *Haversine* untuk memperkuat konteks spasial rekomendasi. Pembangunan sistem dilakukan dengan metodologi Agile (Scrum) agar proses pengembangan bersifat iteratif dan adaptif terhadap kebutuhan pengguna. Hasil evaluasi menunjukkan performa sistem yang baik dengan nilai *Precision* sebesar 0,63, *Recall* sebesar 0,90, dan *F1-Score* sebesar 0,73. Temuan ini menegaskan bahwa integrasi pendekatan berbasis konten dan analisis spasial mampu meningkatkan akurasi rekomendasi wisata yang sesuai dengan preferensi serta lokasi pengguna.

Kata Kunci: Sistem Rekomendasi, *Content-Based Filtering*, Geospasial, *Cosine Similarity*, Agile (Scrum).

ABSTRACT

This study develops a tourism destination recommendation system based on *Content-Based Filtering* integrated with geospatial feature analysis to enhance the relevance and contextual accuracy of recommendations for users. The system addresses the limitations of existing tourism recommendation platforms that primarily focus on popularity and ratings without considering users' location proximity and personal preferences. The dataset used in this research originates from Tourism in Indonesia (Kaggle), focusing on the Jakarta and Bandung regions. Text features are extracted using the *Term Frequency-Inverse Document Frequency* (TF-IDF) method, while the similarity between destinations is measured using *Cosine Similarity*. Additionally, geographic distances are analyzed through the *Haversine* formula to strengthen the spatial context of the recommendations. The system was developed using the Agile (Scrum) methodology to ensure an iterative and adaptive development process aligned with user needs. Evaluation results indicate strong system performance, achieving a *Precision* of 0.63, *Recall* of 0.90, and an *F1-Score* of 0.73. These findings demonstrate that integrating content-based and spatial analysis approaches effectively improves the accuracy and personalization of tourism recommendations based on users' preferences and location context.

Keywords: Recommendation System, *Content-Based Filtering*, Geospatial, *Cosine Similarity*, Agile (Scrum)

1. PENDAHULUAN

Pariwisata merupakan salah satu sektor penting yang berperan besar terhadap perekonomian di Indonesia. Pertumbuhan jumlah wisatawan domestik dan internasional dari tahun ke tahun, ditambah dengan semakin banyak nya destinasi baru yang dikembangkan pemerintah daerah maupun swasta, menunjukkan pesatnya perkembangan industri di bidang ini. Namun perkembangan yang pesat ini juga menimbulkan tantangan bagi wisatawan dalam memilih destinasi yang sesuai dengan kebutuhan nya. Banyaknya destinasi yang tersebar luas serta beragam karakteristik wisatawannya, seperti jenis destinasi yang disukai, rekan perjalanan nya, dan kondisi lingkungan nya, membuat proses rekomendasi menjadi semakin kompleks. Di sisi lain, sebagian besar sistem rekomendasi wisata yang ada saat ini masih terbatas pada popularitas atau rating tanpa mempertimbangkan konteks lokasi dan kebutuhan personal pengguna nya. Hal ini menuntut adanya inovasi atau pembaharuan dalam pengembangan sistem rekomendasi yang lebih adaptif, dengan mengintegrasikan preferensi pengguna dan faktor geospasial untuk menghasilkan rekomendasi destinasi wisata yang lebih relevan dan sesuai kebutuhan pengguna nya.

Perkembangan teknologi digital, khususnya dalam bidang *machine learning* dan sistem rekomendasi, telah membawa dampak besar dalam sektor pariwisata di Indonesia. Kegiatan wisata merupakan salah satu pendorong ekonomi. Didukung dengan geliat investasi baru dan pengembangan infrastruktur, mendorong pemerintah dan pelaku industri untuk berinovasi dan memperkenalkan berbagai destinasi baru yang menarik. Akibatnya, dalam beberapa tahun terakhir, Indonesia mengalami peningkatan yang signifikan dalam jumlah objek wisata komersial [1]. Berdasarkan data Badan Pusat Statistik [1], jumlah perjalanan domestik pada tahun 2023 mencapai lebih dari 734 juta perjalanan, dimana angka ini telah meningkat 12,3% dibandingkan dengan tahun sebelumnya. Perkembangan ini tidak lepas dari kemudahan akses transportasi, promosi secara daring, serta kemunculan platform yang telah mengintegrasikan teknologi *machine learning* dalam merekomendasikan destinasi dan akomodasi kepada pengguna.

Pertumbuhan ini menunjukkan potensi besar dalam industri pariwisata indonesia. Namun, seiring berkembangnya aktivitas pariwisata, preferensi wisatawan juga kini semakin kompleks. Selain mempertimbangkan faktor harga, rating, dan lokasi, wisatawan juga menginginkan rekomendasi yang sesuai dengan kebutuhan spesifik mereka, seperti jenis teman liburan (sendiri, keluarga, teman), jenis destinasi (alam, kuliner, sejarah, petualangan), serta preferensi suhu di lokasi wisata (dingin atau panas). Sistem rekomendasi yang hanya berbasis konten dan geospasial saat ini masih belum memadai untuk mengakomodasi kompleksitas tersebut.

Beberapa permasalahan yang dapat diidentifikasi yaitu, sistem rekomendasi wisata yang ada belum mengintegrasikan faktor lokasi geografis(jarak dan kota pengguna) secara optimal dengan preferensi pengguna, serta belum memiliki fitur filtering dinamis untuk memilih kategori rekan liburan (keluarga, teman, sendiri), tipe destinasi(alam, kuliner, sejarah), maupun preferensi suhu (hangat, sejuk). Selain itu, pengembangan sistem sebelumnya belum menerapkan metode Agile (Scrum). Preferensi wisata yang semakin beragam tidak hanya harga dan rating, tetapi juga jenis wisata dan tingkat keramaian belum sepenuhnya diadaptasi oleh sistem rekomendasi yang ada.

Penelitian ini merancang sebuah sistem penunjang destinasi wisata yang berlandaskan *Content-Based Filtering* dengan memanfaatkan dataset Destinasi Pariwisata Indonesia dari Kaggle, yang terfokus pada tempat-tempat di Jakarta dan Bandung. Teknik yang diterapkan mencakup *Cosine Similarity* dan analisis aspek geospasial seperti lokasi, kota, serta jarak dari pengguna. Preferensi pengguna mencakup rekan wisata, tipe destinasi, serta kondisi lingkungan, yang dikumpulkan melalui masukan yang dilakukan secara manual. Sistem ini dikembangkan sebagai aplikasi web yang interaktif menggunakan metodologi Agile (Scrum), dan dievaluasi berdasarkan akurasi rekomendasi seperti *precision* dan *recall*, beserta kepuasan pengguna dalam rentang yang terbatas, yang dikumpulkan melalui masukan yang dilakukan secara manual.

Untuk itu, sistem rekomendasi yang dikembangkan pada penelitian ini menggunakan metode *Content-Based Filtering* dengan algoritma *K-Nearest Neighbors* (KNN) dan pengukuran kemiripan menggunakan *Cosine Similarity*, yang memadukan fitur konten destinasi dan fitur geospasial serta preferensi pengguna. Sistem yang dikembangkan ini diharapkan mampu memberikan rekomendasi wisata yang relevan dan personal sesuai kebutuhan pengguna. Penelitian ini dapat menjadi dasar bagi pengembangan sistem rekomendasi wisata yang lebih adaptif, dengan mengusulkan sistem rekomendasi

destinasi wisata berbasis *Content-Based Filtering* dengan menggabungkan konteks geospasial dan preferensi pengguna secara bersamaan agar rekomendasi yang dihasilkan lebih dinamis dan akurat.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian terdahulu seperti pada penelitian Penerapan Metode *Content-Based Filtering* dalam rekomendasi sistem wisata mengidentifikasi beberapa kendala utama dalam pengembangan sistem rekomendasi wisata [2]. Antara lain, kurangnya integrasi data lokasi dan preferensi *personal* secara menyeluruh menyebabkan rekomendasi yang dihasilkan kurang kontekstual dan praktis, masalah *sparsity* atau kelangkaan data rating pengguna muncul dalam sistem *collaborative filtering* yang mengakibatkan penurunan akurasi rekomendasi serta kualitas data terutama deskripsi destinasi yang tidak terstruktur juga mempengaruhi performa sistem rekomendasi. Untuk mengatasi permasalahan tersebut, metode *Content-Based Filtering* menjadi solusi utama karena mampu memberikan rekomendasi berdasarkan kemiripan atribut konten destinasi dan preferensi pengguna, dengan pembobotan kata-kata pada deskripsi menggunakan TF-IDF dan pengukuran kemiripan dengan *Cosine Similarity* [2]. TF-IDF dipilih karena kemampuan untuk menemukan kata-kata yang memiliki bobot besar dalam kalimat atau paragraf, yang penting untuk menemukan kata-kata yang memiliki bobot besar dalam kalimat atau paragraf, yang penting untuk menemukan kata beracun [3]. Selain itu, penerapan *hybrid filtering* yang menggabungkan *content-based* dan *collaborative filtering* dapat meningkatkan akurasi dan mengatasi masalah *sparsity* data [4].

Studi lainnya seperti dalam Kusuma Adi Achmad [5] yang berjudul “Model Rekomendasi Destinasi Pariwisata Berbasis Konteks Sosio-Pengguna” mengembangkan sistem rekomendasi dengan pendekatan *context-aware recommendation* yang mempertimbangkan faktor sosial dan karakteristik pengguna. Hasil penelitian menunjukkan bahwa pendekatan berbasis konteks sosio-pengguna mampu memberikan rekomendasi yang lebih sesuai dengan preferensi individu dibandingkan sistem konvensional, meskipun masih memiliki keterbatasan dalam integrasi aspek spasial dan konteks waktu nyata.

2.2 Aplikasi Sistem Rekomendasi Wisata dan Metode Content-Based Filtering

Application yang artinya penerapan, lamaran, penggunaan. Secara istilah aplikasi adalah program siap pakai yang direka untuk melaksanakan suatu fungsi bagi pengguna atau aplikasi yang lain dan dapat digunakan oleh sasaran yang dituju [6]. Metode *Content-Based Filtering* bekerja dengan melihat kemiripan item baru dengan item yang sebelumnya [7].

2.3 K-Nearest Neighbors (KNN) dan Cosine Similarity

K-Nearest Neighbors (KNN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN [8]. KNN diaplikasikan untuk mengukur tingkat kesamaan antara tempat wisata berdasarkan atribut konten serta fitur geospasialnya. Metode *Cosine Similarity* merupakan metode untuk menghitung kesamaan antara dua buah objek yang dinyatakan dalam dua buah vector dengan menggunakan *keywords* dari sebuah dokumen sebagai ukuran [9]. *Cosine Similarity* berfungsi sebagai ukuran kemiripan yang digunakan oleh algoritma KNN dalam menentukan k-nearest neighbors.

2.4 Integrasi Preferensi Pengguna dan Faktor Geospasial

Integrasi antara preferensi pengguna dan faktor geospasial telah menjadi elemen kunci dalam pengembangan sistem rekomendasi destinasi wisata yang efektif. Hal ini berarti menggabungkan model preferensi individu dengan informasi lokasi, jarak ke pengguna, dan koordinat geografis, sehingga rekomendasi yang diberikan dapat mempertimbangkan baik apa yang diinginkan oleh pengguna [10]. Preferensi pengguna sebagai profil dinamis yang harus dimodelkan bersama konteks spasial untuk menghasilkan rekomendasi yang benar-benar kontekstual dan personal [11].

3. METODE PENELITIAN

3.1 Tahapan Penelitian

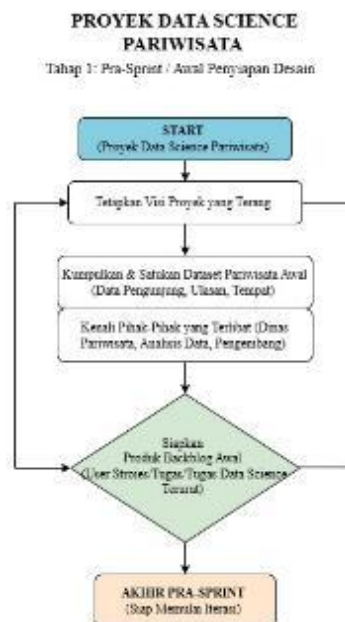
Tahapan penelitian terdiri atas:

1. Pengumpulan Data
2. Pra-pemrosesan Data
3. Ekstraksi Fitur (TF-IDF)
4. Perhitungan Kemiripan (Cosine Similarity)
5. Analisis Geospasial (Haversine)
6. Evaluasi Kinerja Sistem

Metode *Agile* digunakan dalam penelitian ini karena dapat memberikan fleksibilitas dan kecepatan dalam pengembangan sistem rekomendasi. Metode ini digunakan secara bertahap melalui serangkaian *sprint* yang berulang, sehingga hasil pengembangannya dapat langsung dinilai dan disesuaikan dengan kebutuhan pengguna. *Scrum* dipilih sebagai kerangka kerja *Agile* yang digunakan. Menurut Panduan *Scrum* dan literatur praktiknya, *Scrum* adalah sebuah *framework* yang menekankan pengembangan iteratif melalui *sprint* pendek, keterlibatan aktif pihak yang berkepentingan, serta peran dan artefak yang terdefinisi untuk mendukung peningkatan berkelanjutan [12].

Dalam penelitian ini, kerangka kerja *Scrum* diimplementasikan untuk mengatur proses kerja tim secara kolaboratif. Tahapan utama yang dilaksanakan meliputi: *Sprint Planning* yaitu penyusunan tujuan *sprint* dan pemilihan item *Backlog* yang akan dikerjakan; Pelaksanaan *Sprint* dengan pelaksanaan tugas harian (*daily scrum*); Penilaian hasil di akhir *sprint* (*sprint review* dan *retrospective*). Dengan proses tersebut pengembangan sistem dapat berlangsung secara adaptif, dengan fokus pada pengiriman fitur yang bernilai tinggi dan peningkatan berkelanjutan berdasarkan tanggapan pengguna.

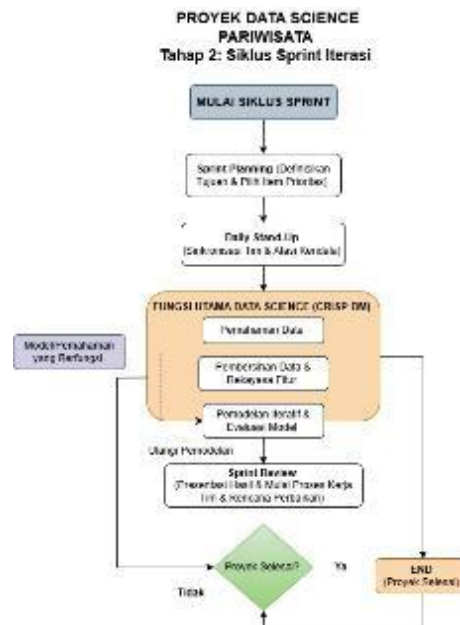
Sistem rekomendasi ini menggunakan algoritma *K-Nearest Neighbors* (KNN). Algoritma ini dipilih karena prinsip kerjanya yang sederhana namun efisien dalam menilai kesamaan antar data. KNN diaplikasikan untuk mengukur tingkat kesamaan antara tempat wisata berdasarkan atribut konten serta fitur geospasialnya. Saat pengguna menginput preferensi atau melakukan pencarian maka sistem ini akan mencari sejumlah destinasi yang sangat mendekati dengan input yang dimasukkan oleh pengguna tersebut. Hasilnya kemudian ditampilkan sebagai rekomendasi yang paling relevan dengan kebutuhan pengguna secara urut.



Gambar.1 Tahapan Scrum

A. Pre-Sprint / Awal Penyiapan Desain

- Menetapkan visi proyek yang terang.
- Mengumpulkan dan menyatukan dataset pariwisata awal (seperti data pengunjung, ulasan, tempat wisata) yang akan dimanfaatkan.
- Mengenali pihak-pihak yang terlibat, seperti Dinas Pariwisata, analis data, serta pengembang.
- Menyiapkan *Product Backlog* awal yang mencakup *User Stories* atau tugas-tugas utama *data science* yang diurutkan berdasarkan nilai bisnis, seperti pembersihan data, analisis *trend* musiman, dan pengembangan model rekomendasi destinasi.



Gambar 2. Tahapan Scrum

B. Siklus *Sprint* Iterasi

- Sprint Planning*: Definisikan tujuan *sprint* dan pilih item dengan prioritas tertinggi dari *Product Backlog* yang akan dikerjakan dalam *sprint* itu.
- Daily Stand-up*: Adakan pertemuan singkat setiap hari untuk menyelaraskan langkah tim dan mengatasi kendala, seperti data yang tidak berkualitas atau masalah performa model.
- Fungsi Utama *Data Science*: Setiap *sprint*, lakukan siklus berulang dalam proses pemahaman data, pembersihan data, rekayasa fitur, pemodelan yang bersifat iteratif, dan evaluasi model berdasarkan pendekatan CRISP-DM. Tujuannya adalah menciptakan model atau pemahaman yang berfungsi.
- Sprint Review*: Tampilkan hasil kerja *sprint*, seperti *dashboard* analisis tren atau model rekomendasi dasar kepada pemangku kepentingan, untuk memperoleh umpan balik secara langsung.
- Sprint Retrospective*: Evaluasi tim mengenai proses kerja *sprint* untuk merencanakan perbaikan yang akan diimplementasikan di *sprint* selanjutnya.

3.2 Rumus yang digunakan

Cosine Similarity:

$$\text{Sim}(A, B) = \frac{\sum A_i B_i}{\sqrt{\sum A_i^2} \times \sqrt{\sum B_i^2}} \dots \dots \dots (1)$$

Haversine Distance:

$$d = 2r \times \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta \text{lat}}{2} \right) + \cos(\text{lat}_1) \cos(\text{lat}_2) \sin^2 \left(\frac{\Delta \text{long}}{2} \right)} \right) \dots \dots \dots (2)$$

3.3 Rancangan Sistem

Sistem terdiri dari beberapa komponen utama, yaitu :

Input: data pengguna dan destinasi wisata

Proses: ekstraksi fitur TF-IDF dan analisis lokasi

Output: rekomendasi destinasi berdasarkan kemiripan konten dan lokasi

3.4 Use Case Diagram

Use case diagram digunakan untuk menggambarkan interaksi antara pengguna dan admin dengan sistem rekomendasi. Diagram ini menunjukkan fungsi-fungsi yang dapat dilakukan oleh pengguna dan sistem. Berdasarkan hasil analisis kebutuhan, terdapat dua peran utama yaitu User dan Admin. User dapat melakukan registrasi, login, melihat daftar destinasi, detail destinasi, memberi rating dan mendapatkan rekomendasi. Sementara Admin dapat melakukan Kelola destinasi wisata, Kelola user, juga monitoring hasil rekomendasi.



Gambar 3. Use Case Diagram

4. PEMBAHASAN

4.1 Dataset

Tabel 1. Dataset

No	Place_Name	Description	Category	City	Partner	Temperature
1	Monumen Nasional	Monumen nasional	Budaya	Jakarta	Keluarga	Hangat
2	Kota Tua	Kota Tua di Jakarta yang....	Budaya	Jakarta	Teman	Hangat
3	Dunia Fantasi	Dunia fantasia atau juga....	Tempat Hiburan	Jakarta	Keluarga	Hangat
4	Taman Mini	Merupakan suatu....	Taman Hiburan	Jakarta	Keluarga	Hangat
5	Atlantis	Atau dikenal....	Taman Hiburan	Jakarta	Keluarga	Hangat

.....						
241	Bandros City Tour	Adalah bus wisata yang...	Budaya	Bandung	Teman	Sejuk
242	Kyouto Floating Market	Sebuah wisata kekinian	Budaya	Bandung	Teman	Sejuk
243	Rainbow Garden	Salah satu taman rekreasi	Cagar Alam	Bandung	Teman	Sejuk
244	Kota Mini	Destinasi menarik...	Taman Hiburan	Bandung	Keluarga	Sejuk
245	Chingu Cafe Little Soul	Selain populer karena...	Taman Hiburan	Bandung	Teman	Sejuk

4.2 Hasil Pengolahan Data

Sumber data yang digunakan dalam penelitian ini adalah dataset publik *Tourism in Indonesia* yang diunduh dari platform *Kaggle*. Dataset ini terdiri atas empat tabel utama, yaitu *user.csv*, *tourism_with_id.csv*, *tourism_rating.csv*, dan *package_tourism.csv*. Data tersebut kemudian diolah untuk menghasilkan model sistem rekomendasi destinasi wisata. Tahapan pengolahan data diawali dengan pembersihan data (data cleaning) terhadap atribut yang kosong, serta penyamaan format teks pada deskripsi destinasi wisata. Proses selanjutnya adalah transformasi data rating menjadi nilai biner, di mana nilai *rating* ≥ 4 dikategorikan sebagai “disukai” dan *rating* < 4 dikategorikan sebagai “tidak disukai”. Hasil preprocessing menunjukkan bahwa sebagian besar pengguna memberikan rating tinggi terhadap destinasi populer seperti Monas, Taman Mini Indonesia Indah, Farm House Lembang, dan Dufan.

4.3 Hasil Ekstraksi Fitur dan Pembobotan

Proses ekstraksi fitur dilakukan pada kolom deskripsi tempat wisata menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk mendapatkan representasi numerik dari kata kunci penting yang menggambarkan karakteristik setiap destinasi wisata. Kata-kata yang sering muncul dalam deskripsi, seperti “keluarga”, “alam”, “edukasi”, dan “sejarah”, menjadi fitur penting yang mewakili kategori wisata tersebut.

Hasil pembobotan TF-IDF kemudian digunakan untuk menghitung tingkat kemiripan antar destinasi menggunakan rumus Cosine Similarity. Nilai kemiripan berkisar antara 0 dan 1; semakin mendekati 1 berarti dua destinasi memiliki deskripsi konten yang semakin mirip. Misalnya, destinasi “Farm House Lembang” memiliki tingkat kemiripan tinggi dengan “Floating Market Lembang” karena keduanya sama-sama menawarkan konsep wisata keluarga dan alam.

4.4 Implementasi Sistem Rekomendasi

Sistem rekomendasi ini menerapkan *Content-Based Filtering* yang menggabungkan dua pendekatan utama:

1. Kesamaan Konten (Content Similarity) yang dihitung berdasarkan kemiripan deskripsi wisata.
2. Analisis Geospasial yang memperhitungkan jarak antar destinasi menggunakan rumus *Haversine*.

Bobot total rekomendasi dihitung dengan persamaan (3):

$$\text{Score}(i) = \alpha \times \text{Sim}_{\text{content}}(i) + \beta \times \text{Sim}_{\text{geo}}(i) \dots \dots \dots (3)$$

dengan α dan β merupakan bobot penyesuaian. Dalam penelitian ini digunakan nilai $\alpha = 0.7$ dan $\beta = 0.3$, sehingga sistem lebih mengutamakan kesamaan konten namun tetap memperhatikan faktor kedekatan geografis.

Proses implementasi dilakukan menggunakan *RapidMiner Studio 2025* sebagai alat bantu dalam pemodelan dan evaluasi sistem. Tahapan implementasi terdiri atas:

- a. *Import Dataset* dari Kaggle ke RapidMiner.

- b. *Preprocessing* menggunakan operator *Replace Missing Values* dan *Nominal to Text*.
- c. *Text Processing* dengan operator *Process Documents from Data* dan *TF-IDF*.
- d. *Similarity Computation* dengan operator *Cosine Similarity*.
- e. *Evaluation* menggunakan *Performance (Classification)* untuk menghitung metrik *Precision*, *Recall*, dan *F1-Score*.

4.5 Hasil Evaluasi Sistem

Evaluasi kinerja sistem dilakukan terhadap lima pengguna acak, yaitu U1004, U1021, U1053, U1072, dan U1105, yang dipilih dari dataset *user.csv*. Setiap pengguna diberikan hasil rekomendasi sebanyak lima destinasi wisata berdasarkan profil preferensi mereka. Evaluasi dilakukan dengan membandingkan hasil rekomendasi sistem terhadap data aktual destinasi yang disukai oleh pengguna (berdasarkan rating ≥ 4).

Perhitungan dilakukan menggunakan rumus berikut:

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN}, \text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Hasil evaluasi ditampilkan pada Tabel 1.

Tabel 2. Hasil Evaluasi Sistem Rekomendasi Wisata

User_Id	TP	FP	FN	Precision	Recall	F1-Score
U1004	3	2	1	0.60	0.75	0.67
U1021	3	2	0	0.60	1.00	0.75
U1053	3	2	0	0.60	1.00	0.75
U1072	3	1	1	0.75	0.75	0.75
U1105	3	2	0	0.60	1.00	0.75
Rata-rata	3.0	1.8	0.4	0.63	0.90	0.73

Dari tabel tersebut, sistem menunjukkan rata-rata *Precision* sebesar 0.63, *Recall* sebesar 0.90, dan *F1-Score* sebesar 0.73. Nilai *Precision* mengindikasikan bahwa sekitar 63% dari hasil rekomendasi merupakan destinasi yang benar-benar disukai pengguna, sedangkan nilai *Recall* menunjukkan bahwa 90% dari destinasi yang disukai pengguna berhasil direkomendasikan oleh sistem.

4.6 Tampilan Antarmuka

A. Halaman Login

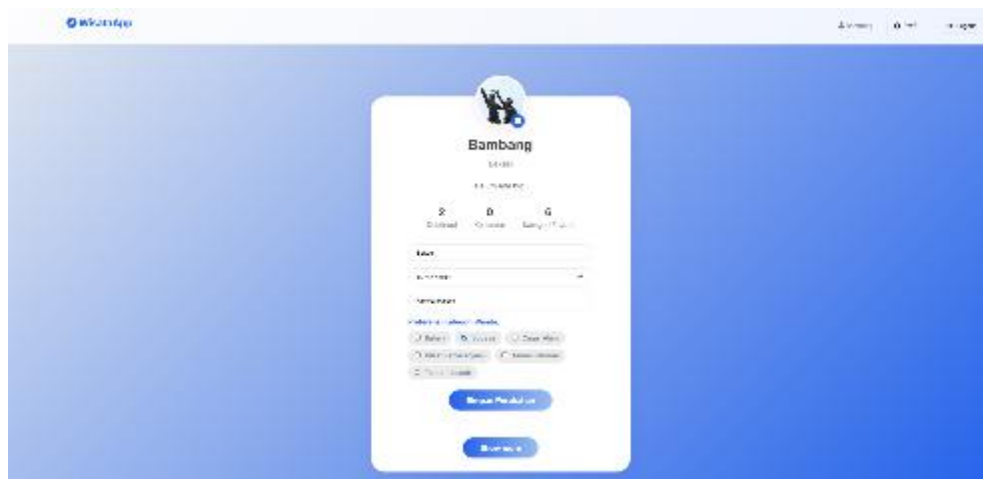
Pengguna dapat login ke akun mereka jika sudah terdaftar atau register (membuat akun baru) jika belum mempunyai akun.



Gambar 4. Halaman *Login*

B. Halaman Profil

Pada Halaman profil, terdapat fitur dimana pengguna dapat memperbarui data pribadi mulai dari foto profil, Alamat, No. Telepon dan data diri lain.



Gambar 5. Halaman Profil

C. Beranda

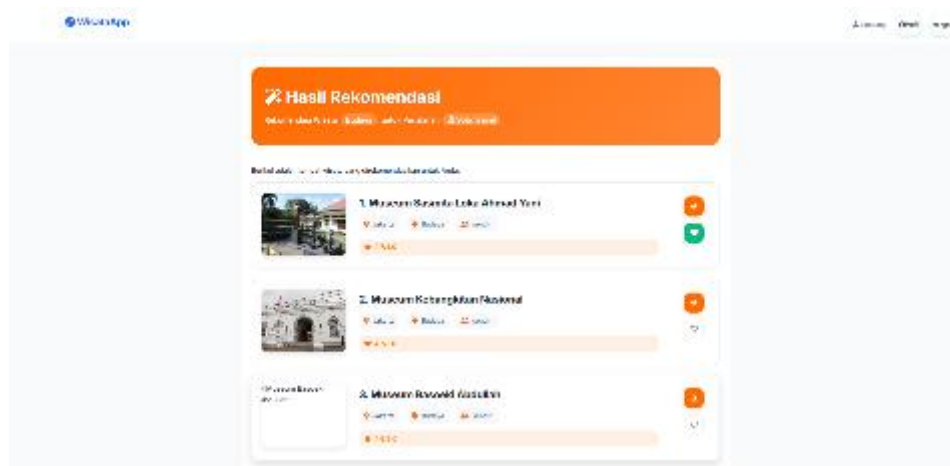
Ini merupakan halaman utama pada website, dimana user dapat mencari rekomendasi sesuai dengan preferensi penggunaanya.



Gambar 6. Beranda

D. Hasil Rekomendasi

Menampilkan halaman hasil rekomendasi yang didapat setelah memasukkan preferensi pada halaman pertama.



Gambar 7. Hasil Rekomendasi

5. KESIMPULAN

Sistem rekomendasi destinasi wisata berbasis Content-Based Filtering dan analisis fitur geospasial menggunakan dataset publik Tourism in Indonesia telah berhasil dikembangkan. Sistem dirancang untuk menampilkan rekomendasi wisata yang relevan berdasarkan kemiripan onten deskripsi dan jarak geografis antar lokasi. Dari hasil evaluasi nilai rata-rata Precision sebesar 0.63, Recall sebesar 0.90, dan F1-Score sebesar 0.73 telah menunjukkan bahwa model memiliki kinerja yang cukup baik. Hal ini menunjukkan bahwa metode yang diterapkan berhasil dalam memberikan rekomendasi yang tepat dan relevan berdasarkan referensi pengguna.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Data yang digunakan masih terbatas pada wilayah Jakarta dan Bandung, serta fitur yang dianalisis hanya mencakup deskripsi dan rating pengguna. Penelitian selanjutnya dapat memperluas cakupan wilayah dan menambahkan fitur lain seperti ulasan pengguna, kategori wisata, dan harga tiket untuk meningkatkan akurasi serta keberagaman rekomendasi. Selain itu, pengembangan sistem ke arah model *hybrid filtering* disarankan agar sistem mampu mengatasi permasalahan *cold-start* dan memberikan rekomendasi yang lebih adaptif.

DAFTAR PUSTAKA

- [1] BPS, “Statistik Objek Daya Tarik Wisata Tahun 2023,” 2024, [Online]. Available: <https://www.bps.go.id/publication/2024/12/20/ebca1af66c31bbdf2b912be9/statistik-objek-daya-tarik-wisata-2023.html>
- [2] A. Sahgal, “Опыт аудита обеспечения качества и безопасности медицинской деятельности в медицинской организации по разделу «Эпидемиологическая безопасность» No Title,” *Вестник Росздравнадзора*, vol. 4, no. 1, pp. 9–15, 2024.
- [3] R. Budiarti, A. I. Purnamasari, A. Bahtiar, and E. Wahyudin, “Peningkatan Model Prediksi Konten Cyberbullying Pada Media Sosial Instagram Menggunakan Algoritma Support Vector Machine,” *J. Inf. Interaktif*, vol. 10, no. 1, pp. 63–73, 2025.
- [4] Y. I. Lubis, D. J. Napitupulu, and A. S. Dharma, “Implementasi Metode Hybrid Filtering (Collaborative dan Content-based) untuk Sistem Rekomendasi Pariwisata,” *12th Conf. Inf. Technol. Electr. Eng.*, pp. 6–8, 2020.
- [5] “Model Rekomendasi Destinasi Pariwisata Berbasis Konteks Sosio-Pengguna KUSUMA ADI ACHMAD, Ir. Lukito Edi Nugroho, M. Sc., Ph. D.; Prof. Ir. Achmad Djunaedi, MUP., Ph. D.; Widyawan,” vol. 0, p. 2019, 2019.
- [6] F. Muhammad, N. R. Fahrudin, J. K. No, K. Kesambi, and K. Cirebon, “PERANCANGAN UI/UX APLIKASI KESEHATAN PATIENT-CENTRIC MENGGUNAKAN METODE USER CENTERED DESIGN BERBASIS MOBILE (STUDI KASUS: KLINIK MUHAMMADIYAH),” vol. 10, no. 3, pp. 217–226, 2025.
- [7] R. F. Oeyliawan and D. Gunawan, “Aplikasi Rekomendasi Buku Pada Katalog Perpustakaan Universitas Multimedia Nusantara Menggunakan Vector Space Model,” *Ultim. J. Tek. Inform.*, vol. 9, no. 2, pp. 97–105, 2017, doi: 10.31937/ti.v9i2.639.
- [8] R. N. F. U. S. Rut Samuel, “[2]+JBIDAI+Vol.4,+No.1,+Juni+2018+Hal+9-14,” *Big Data Anal. Artif. Intell.*, vol. 1, no. 1, pp. 1–6, 2021.
- [9] R. T. Wahyuni, D. Prastiyanto, and E. Suprptono, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi,” vol. 9, no. 1, 2017.
- [10] L. Ravi and S. Vairavasundaram, “A Collaborative Location Based Travel Recommendation System through Enhanced Rating Prediction for the Group of Users,” *Comput. Intell. Neurosci.*, vol. 2016, 2016, doi: 10.1155/2016/1291358.
- [11] J. Borràs, A. Moreno, and A. Valls, “Intelligent tourism recommender systems: A survey,” *Expert Syst. Appl.*, vol. 41, no. 16, pp. 7370–7389, 2014, doi: 10.1016/j.eswa.2014.06.007.
- [12] K. Schwaber and J. Sutherland, “The Scrum Guide: The Definitive The Rules of the Game,” *Scrum.Org and ScrumInc*, no. November, p. 19, 2017, [Online]. Available: <http://www.scrumguides.org/docs/scrumguide/v1/Scrum-Guide-US.pdf>