

**SISTEM REKOMENDASI TUJUAN WISATA DI JAWA TIMUR BERBASIS
CONTENT-BASED FILTERING MENGGUNAKAN *COSINE SIMILARITY***

SKRIPSI

Oleh:
HILYA ZAHRA ALIFA
NIM. 220605110171



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**SISTEM REKOMENDASI TUJUAN WISATA DI JAWA TIMUR
BERBASIS *CONTENT-BASED FILTERING* MENGGUNAKAN
*COSINE SIMILARITY***

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:
HILYA ZAHRA ALIFA
NIM. 220605110171

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

HALAMAN PERSETUJUAN

SISTEM REKOMENDASI TUJUAN WISATA DI JAWA TIMUR BERBASIS *CONTENT-BASED FILTERING* MENGGUNAKAN *COSINE SIMILARITY*

SKRIPSI

Oleh:
HILYA ZAHRA ALIFA
NIM. 220605110171

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 12 Juni 2026

Pembimbing I,



Syahiduz Zaman, M.Kom
NIP. 19700502 200501 1 005

Pembimbing II,



Ashri Shabrina Afrah, M.T
NIP. 19900430 202012 2 003

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Syahrudin, M.Kom
NIP. 1981081010 201903 1 012

HALAMAN PENGESAHAN

SISTEM REKOMENDASI TUJUAN WISATA DI JAWA TIMUR BERBASIS *CONTENT-BASED FILTERING* MENGGUNAKAN *COSINE SIMILARITY*

SKRIPSI

Oleh:

HILYA ZAHRA ALIFA

NIM. 220605110171

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 19 Juni 2026

Susunan Dewan Penguji

Ketua Penguji	: <u>Dr. Totok Chamidy, M.Kom</u> NIP. 19691222 200604 1 001
Anggota Penguji I	: <u>Fajar Rohman Hariri, M.Kom</u> NIP. 19890515 201801 1 001
Anggota Penguji II	: <u>Syahiduz Zaman, M.Kom</u> NIP. 19700502 200501 1 005
Anggota Penguji III	: <u>Ashri Shabrina Afrah, M.T</u> NIP. 19900430 202012 2 003

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Hilya Zahra Alifa
NIM : 220605110171
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Sistem Rekomendasi Tujuan Wisata di Jawa Timur
Berbasis *Content-Based Filtering* Menggunakan
Cosine Similarity

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 25 Juni 2026

Yang membuat pernyataan,



Hilya Zahra Alifa
NIM. 220605110171

MOTTO

*“Maka, sesungguhnya beserta kesulitan ada kemudahan”
(QS. Al-Insyirah:5)*

*“Life can be heavy, especially if you try to carry it all at once.
Part of growing up and moving into new chapters of your life is
about catch and release”
-Taylor Swift*

HALAMAN PERSEMBAHAN

Alhamdulillah *rabbil 'alamin*, segala puji syukur penulis panjatkan ke hadirat Allah *Subhanahu Wa Ta'ala* yang telah melimpahkan rahmat, karunia, dan hidayah-Nya, sehingga penulis dapat menuntaskan skripsi ini dengan baik.

Dengan penuh rasa syukur, penulis persembahkan kepada:

Kedua Orang Tua Penulis,

Yang tak pernah lelah mendoakan dan mendukung kelancaran studi penulis hingga titik ini. Semoga setiap kebaikan yang telah diberikan menjadi amal jariyah yang terus mengalir dan dibalas berlipat oleh Allah *Subhānahu wa Ta'ālā*.

Keluarga Penulis,

yang selalu hadir dengan kehangatan dan dukungan di saat yang paling dibutuhkan. Kata-kata kalian yang sederhana sering kali menjadi cukup untuk membuat penulis kembali bangkit dan melanjutkan langkah. Terima kasih telah menjadi pengingat bahwa penulis tidak pernah benar-benar sendiri.

Untuk Diri Penulis,

Terima kasih sudah bertahan dan terus berjalan melewati segala tantangan yang semesta hadirkan.

KATA PENGANTAR

Assalamu 'alaikum Wr. Wb.

Alhamdulillah, segala puji dan rasa syukur penulis panjatkan ke hadirat Allah *Subhanahu Wa Ta'ala* atas rahmat, karunia, dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Sistem Rekomendasi Tujuan Wisata di Jawa Timur Berbasis *Content-Based Filtering* Menggunakan *Cosine Similarity*”. Sholawat serta salam tetap tercurahkan kepada junjungan kita Nabi Muhammad *Shallallahu Alaihi Wasallam*, yang menjadi suri tauladan dan semoga kita mendapat syafaat di akhir kelak, Aamiin.

Dalam penyusunan skripsi ini, penulis menyadari bahwa banyak pihak telah memberikan doa, bantuan, arahan, dan dukungan. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Agus Mulyono, M.Kes., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Supriyono, M.Kom., selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Syahiduz Zaman, M.Kom., selaku Dosen Pembimbing I yang telah meluangkan banyak waktu untuk memberikan bimbingan, arahan, dan masukan yang berharga selama penyusunan skripsi ini.
5. Ashri Shabrina Afrah, M.T., selaku Dosen Pembimbing II yang telah sabar dalam memberikan ilmu dan arahan dalam setiap tahap penulisan skripsi ini.

6. Dr. Totok Chamidy, M.Kom., selaku Ketua Penguji atas koreksi, kritik, dan saran yang diberikan sehingga sangat membantu dalam penyempurnaan skripsi ini.
7. Fajar Rohman Hariri, M.Kom., selaku Anggota Penguji I atas segala saran dan masukan yang memperkuat kualitas skripsi ini sejak seminar proposal hingga sidang skripsi.
8. Seluruh dosen, tenaga kependidikan, laboran, serta jajaran staf Program Studi Teknik Informatika yang telah memberikan dukungan selama penulis menempuh studi.
9. Kedua orang tua tercinta, yang selama ini menjadi alasan penulis untuk terus bertahan dan tidak menyerah. Terima kasih atas doa yang tak pernah putus, pengorbanan yang tak pernah diperlihatkan, dan kasih sayang yang tidak pernah meminta balasan. Semoga setiap tetes keringat yang dicurahkan untuk penulis menjadi amal jariyah yang terus mengalir.
10. Kakak, Mbah Uti dan Mbah Kakung penulis, yang selalu hadir di saat yang tepat dengan kata-kata yang cukup untuk membuat penulis kembali bangkit dan melanjutkan langkah.
11. Husnul Khatimah, terima kasih telah hadir menjadi teman penulis sejak hari-hari pertama menjadi mahasiswa baru, berbagi satu kamar di ma'had, hingga perkuliahan ini benar-benar selesai. Terima kasih juga atas segala bantuan dan dukungan yang menemani perjalanan penulis dari awal perkuliahan hingga akhir.

12. Laudza Atsila Prasetyo, teman yang menemani baik di bangku kuliah maupun di masa magang. Terima kasih atas ilmu-ilmu baru yang dibagikan dan kebersamaan yang membuat hari-hari perkuliahan dan kerja praktik terasa lebih ringan dan menyenangkan.
13. Caca dan Silvi, yang hadir di semester akhir namun langsung menjadi bagian penting dari perjalanan ini. Terima kasih sudah berjuang bersama dan saling membantu di fase yang paling melelahkan ini.
14. Khansa, Syifa, Sabhina, Risma Naia, Mayscha, Karina, Aurel, dan Alia, sahabat-sahabat SMA yang meskipun jarak dan kesibukan memisahkan, namun tali silaturahmi tetap terjaga. Terima kasih sudah menjadi bagian dari cerita yang jauh sebelum perkuliahan ini dimulai.
15. Teman-teman seperjuangan angkatan 2022, yang melewati fase yang sama dengan caranya masing-masing. Terima kasih sudah saling menguatkan dan menjaga semangat untuk sampai di titik ini bersama-sama.
16. Seluruh responden dan *annotator* yang telah meluangkan waktu untuk berpartisipasi dalam penelitian ini. Kontribusi kalian sangat berarti dan menjadi bagian penting dari terselesaikannya skripsi ini.
17. Seluruh pihak yang tidak dapat disebutkan satu per satu, namun kehadirannya terasa dan kebaikannya tidak terlupakan. Semoga *Allah Subhānahu wa Ta'ālā* membalas semuanya dengan yang jauh lebih baik.
18. Terakhir, penulis tak henti mengucapkan terima kasih untuk satu orang yang selama ini diam-diam berjuang tanpa henti, penulis sendiri, Hilya Zahra Alifa. Terima kasih sudah bertahan dan terus berjalan melewati segala

tantangan yang semesta hadirkan. Tetap belajar menerima dan mensyukuri, walaupun terkadang yang didapatkan tidak sesuai dengan yang diharapkan. Semoga setiap langkah kecil itu selalu kuat dan dikelilingi oleh orang-orang hebat, serta satu persatu mimpi-mimpi itu akan terjawab.

Malang, 25 Juni 2026

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL	xv
ABSTRAK	xvi
ABSTRACT	xvii
ملخص البحث	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	8
BAB II STUDI PUSTAKA	9
2.1 Penelitian Terdahulu	9
2.2 Sistem Rekomendasi	14
2.3 <i>Content-Based Filtering</i>	15
2.4 <i>Text Preprocessing</i>	17
2.5 <i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	19
2.6 <i>Cosine Similarity</i>	20
2.7 Evaluasi Sistem	21
2.7.1 <i>Precision@K</i>	22
2.7.2 <i>Post-Study System Usability Questionnaire (PSSUQ)</i>	22
BAB III DESAIN DAN IMPLEMENTASI	25
3.1 Desain Penelitian.....	25
3.2 Teknik Pengumpulan Data.....	27
3.2.1 Data Penelitian	28
3.2.1.1 <i>Dataset Deskripsi Destinasi (Konten)</i>	29
3.2.1.2 <i>Dataset Ulasan Pengguna</i>	30
3.3 Desain Sistem.....	31
3.4 Tahap <i>Text Preprocessing</i>	32
3.5 Implementasi Metode.....	36
3.5.1 <i>Content-Based Filtering</i>	36
3.6 Evaluasi Sistem	46
3.6.1 Evaluasi Sistem Rekomendasi	46
3.7 Skenario Uji Coba	47
BAB IV HASIL DAN PEMBAHASAN	49
4.1 Pengumpulan Data	49

4.2	Implementasi <i>Database</i>	50
4.3	Hasil Sistem Rekomendasi.....	51
4.3.1	Hasil <i>Text Preprocessing</i>	52
4.3.2	Perhitungan TF-IDF	53
4.3.3	Perhitungan <i>Cosine Similarity</i>	55
4.3.4	Hasil Rekomendasi.....	57
4.4	Implementasi Sistem	69
4.4.1	Halaman Beranda	69
4.4.2	Halaman Cari Rekomendasi.....	71
4.4.3	Halaman Seluruh Data Destinasi	72
4.4.4	Halaman Rincian Destinasi	73
4.4.5	Halaman Pencarian Rekomendasi.....	76
4.5	Evaluasi Sistem Rekomendasi	78
4.5.1	Evaluasi <i>Precision@K</i>	79
4.5.2	Evaluasi <i>Usability</i> Menggunakan PSSUQ.....	85
4.5.2.1	<i>System Usefulness</i> (SysUse).....	86
4.5.2.2	<i>Information Quality</i> (InfoQual)	86
4.5.2.3	<i>Interface Quality</i> (InterQual)	87
4.5.2.4	<i>Overall</i>	89
4.6	Integrasi Islam.....	90
4.6.1	<i>Mu'ammalah Ma'allah</i> (Hubungan Manusia dengan Allah).....	90
4.6.2	<i>Mu'ammalah Ma'annas</i> (Hubungan Manusia dengan Sesama Manusia).....	91
BAB V KESIMPULAN DAN SARAN		93
5.1	Kesimpulan	93
5.2	Saran.....	94
DAFTAR PUSTAKA		

DAFTAR GAMBAR

Gambar 2.1 Mekanisme Content-Based Filtering.....	16
Gambar 3.1 Diagram Alur Penelitian.....	26
Gambar 3.2 Diagram Desain Sistem.....	31
Gambar 3.3 Diagram Tahapan <i>Text Preprocessing</i>	32
Gambar 4.1 Data Destinasi Wisata	49
Gambar 4.2 Contoh Data Ulasan Pengunjung	50
Gambar 4.3 Implementasi Database Dalam Bentuk ERD	51
Gambar 4.4 Data Destinasi Wisata Setelah Penggabungan Kolom.....	52
Gambar 4.5 Tampilan Halaman Beranda.....	70
Gambar 4.6 Tampilan Halaman Cari Rekomendasi	71
Gambar 4.7 Tampilan Halaman Cari Rekomendasi dengan Hasil Rekomendasi.	72
Gambar 4.8 Tampilan Halaman Seluruh Data Destinasi	73
Gambar 4.9 Tampilan Halaman Rincian Destinasi.....	74
Gambar 4.10 Tampilan Halaman Rincian Destinasi Kesan Pengunjung.....	75
Gambar 4.11 Tampilan Rekomendasi.....	76
Gambar 4.12 Tampilan Halaman Pencarian Rekomendasi (Sebelum Pencarian)	77
Gambar 4.13 Tampilan Halaman Pencarian Rekomendasi (Setelah Pencarian) ..	78

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	12
Tabel 3.1 Sampel Data Deskripsi Destinasi Wisata.....	29
Tabel 3.2 Contoh Sampel Data Ulasan	30
Tabel 3.3 Hasil <i>Text Preprocessing Cleaning</i>	33
Tabel 3.4 Hasil <i>Text Preprocessing Case Folding</i>	34
Tabel 3.5 Hasil <i>Text Processing Tokenizing</i>	34
Tabel 3.6 Hasil <i>Text Preprocessing Normalization</i>	35
Tabel 3.7 Hasil <i>Text Preprocessing Stopword Removal</i>	35
Tabel 3.8 Hasil <i>Text Preprocessing Stemming</i>	36
Tabel 3.9 Contoh Sampel Data Destinasi Wisata	37
Tabel 3.10 Hasil <i>Text Preprocessing</i> Data Destinasi Wisata.....	38
Tabel 3.11 Representasi Dokumen yang Telah Melalui <i>Text Preprocessing</i>	39
Tabel 3.12 Perhitungan TF.....	40
Tabel 3.13 Perhitungan IDF	41
Tabel 3.14 Contoh Sampel Data Deskripsi Setelah Perhitungan TF-IDF	42
Tabel 3.15 Perhitungan <i>Dot Product</i>	43
Tabel 3.16 Perhitungan Panjang Vektor	44
Tabel 3.17 Hasil <i>Cosine Similarity</i>	45
Tabel 3.18 Daftar Pertanyaan PSSUQ	47
Tabel 4.1 Hasil Perhitungan TF	54
Tabel 4.2 Hasil Perhitungan IDF	54
Tabel 4.3 Hasil Perhitungan TF-IDF	55
Tabel 4.4 Perhitungan Cosine Similarity	56
Tabel 4.5 Input Data.....	57
Tabel 4.6 Hasil Rekomendasi	58
Tabel 4.7 Hasil Rekomendasi Skenario 3 (Kueri “Pantai bersih”)	69
Tabel 4.8 Hasil Skenario 1	79
Tabel 4.9 Nilai <i>Precision</i> Skenario 1	80
Tabel 4.10 Hasil Skenario 2	81
Tabel 4.11 Nilai <i>Precision</i> Skenario 2	81
Tabel 4.12 Hasil Skenario 3	82
Tabel 4.13 Nilai <i>Precision</i> Skenario 3	82
Tabel 4.14 Hasil Skenario 4	83
Tabel 4.15 Nilai <i>Precision</i> Skenario 4	84
Tabel 4.16 Nilai Rata-rata <i>Precision</i>	84
Tabel 4.17 Norma Skala PSSUQ	85
Tabel 4.18 Metode PSSUQ Aspek <i>System Usefulness</i>	86
Tabel 4.19 Metode PSSUQ Aspek <i>Information Quality</i>	87
Tabel 4.20 Metode PSSUQ Aspek <i>Interface Quality</i>	87
Tabel 4.21 Metode PSSUQ Aspek <i>Overall</i>	89

ABSTRAK

Alifa, Hilya Zahra. 2026. **Sistem Rekomendasi Tujuan Wisata di Jawa Timur Berbasis *Content-Based Filtering* Menggunakan *Cosine Similarity***. Skripsi. Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Syahiduz Zaman, M.Kom. (II) Ashri Shabrina Afrah, M.T.

Kata kunci: Sistem Rekomendasi, Content-Based Filtering, TF-IDF, Cosine Similarity, Pariwisata Jawa Timur, PSSUQ

Perkembangan sektor pariwisata di Indonesia terus menunjukkan pertumbuhan yang signifikan, tercermin dari peningkatan kunjungan wisatawan asing sebesar 13,01% dan perjalanan wisatawan domestik sebesar 29,72% pada Juli 2025 dibandingkan periode yang sama tahun sebelumnya. Jawa Timur sebagai salah satu provinsi dengan keanekaragaman destinasi wisata memerlukan sistem yang mampu membantu wisatawan menemukan tujuan yang sesuai dengan preferensi mereka secara efisien. Penelitian ini bertujuan membangun sistem rekomendasi destinasi wisata di Jawa Timur menggunakan metode Content-Based Filtering dengan TF-IDF dan Cosine Similarity, serta mengintegrasikan ulasan wisatawan untuk memperkaya representasi konten destinasi. Sistem dikembangkan berbasis web menggunakan Laravel dan FastAPI, mencakup 27 destinasi wisata dengan data ulasan yang bersumber dari berbagai platform media sosial. Evaluasi dilakukan menggunakan dua pendekatan: metrik Precision@K untuk mengukur relevansi rekomendasi, dan Post-Study System Usability Questionnaire (PSSUQ) untuk mengukur kepuasan pengguna terhadap 52 responden. Hasil evaluasi menunjukkan rata-rata Precision@K (K=5) sebesar 0,69 secara keseluruhan, dengan performa tertinggi pada kategori wisata bahari (1,00). Evaluasi PSSUQ menghasilkan skor rata-rata Overall 2,07, System Usefulness 2,02, Information Quality 2,14, dan Interface Quality 2,05, seluruhnya berada di bawah batas bawah norma skala PSSUQ yang menandakan tingkat usability sangat baik. Sistem ini terbukti mampu memberikan rekomendasi yang relevan sesuai preferensi pengguna, meskipun performa dipengaruhi oleh homogenitas kategori destinasi dan panjang kata kunci yang digunakan.

ABSTRACT

Alifa, Hilya Zahra. 2026. **Tourism Destination Recommendation System in East Java Based on Content-Based Filtering Using Cosine Similarity.** Undergraduate Thesis. Department of Informatics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Syahiduz Zaman, M.Kom. (II) Ashri Shabrina Afrah, M.T.

Keywords: *Recommendation System, Content-Based Filtering, TF-IDF, Cosine Similarity, East Java Tourism, PSSUQ*

Tourism in Indonesia has continued to experience significant growth, as reflected by a 13.01% increase in international tourist arrivals and a 29.72% increase in domestic tourist trips in July 2025 compared to the same period in the previous year. As one of Indonesia's provinces with a diverse range of tourist destinations, East Java requires a system capable of helping travelers efficiently identify destinations that match their preferences. This study aims to develop a tourism destination recommendation system for East Java using the Content-Based Filtering method combined with TF-IDF and Cosine Similarity, while integrating tourist reviews to enrich the content representation of each destination. The system was developed as a web-based application using Laravel and FastAPI, covering 27 tourist destinations with review data collected from various social media platforms. The system was evaluated using two approaches: Precision@K to measure recommendation relevance and the Post-Study System Usability Questionnaire (PSSUQ) to assess user satisfaction involving 52 respondents. The evaluation results showed an overall average Precision@5 score of 0.69, with the highest performance achieved in the marine tourism category (1.00). The PSSUQ evaluation produced average scores of 2.07 for Overall, 2.02 for System Usefulness, 2.14 for Information Quality, and 2.05 for Interface Quality. All scores were below the lower bound of the PSSUQ normative scale, indicating an excellent level of system usability. These findings demonstrate that the proposed system is capable of providing relevant recommendations that align with user preferences, although its performance is influenced by the homogeneity of destination categories and the length of the input keywords.

ملخص البحث

عليقة، حلية زهرة. ٢٠٢٦. نظام توصية الوجهات السياحية في جاوة الشرقية القائم على **Content-Based Filtering** باستخدام **Cosine Similarity**. بحث جامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: شاهد الزمن، الماجستير؛ المشرف الثاني: عصري صبرنا أفرح، الماجستير.

الكلمات المفتاحية: نظام التوصية، التصفية القائمة على المحتوى، TF-IDF، تشابه جيب التمام، سياحة جاوة الشرقية، PSSUQ

يشهد قطاع السياحة في إندونيسيا نمواً ملحوظاً، إذ سجل شهر يوليو 2025 ارتفاعاً في عدد الزوار الأجانب بنسبة 13,01% وفي رحلات السياح المحليين بنسبة 29,72% مقارنةً بالفترة ذاتها من العام السابق. وتُعدّ مقاطعة جاوة الشرقية من أكثر المناطق تنوعاً في وجهاتها السياحية، مما يستدعي وجود نظام يُساعد السياح على اكتشاف الوجهات المناسبة لتفضيلاتهم بكفاءة عالية. تهدف هذه الدراسة إلى بناء نظام توصية للوجهات السياحية في جاوة الشرقية باستخدام أسلوب التصفية القائمة على المحتوى (Content-Based Filtering) مع تقنية TF-IDF وخوارزمية تشابه جيب التمام (Cosine Similarity)، مع دمج تقييمات السياح لإثراء تمثيل محتوى الوجهات. طُوّر النظام على شكل تطبيق ويب باستخدام FastAPI و Laravel، ويشمل 27 وجهة سياحية مع بيانات مراجعات مستقاة من منصات التواصل الاجتماعي المختلفة. جرى تقييم النظام وفق منهجين: مقياس $Precision@K$ لقياس دقة التوصيات، واستبيان PSSUQ لقياس رضا المستخدمين على عينة مؤلفة من 52 مستجيباً. أظهرت نتائج التقييم أن متوسط $Precision@K$ (K=5) بلغ 0,69 بشكل إجمالي، مع أعلى أداء في فئة السياحة البحرية (1,00). وأسفر تقييم PSSUQ عن متوسط درجة إجمالية قدرها 2,07، وفائدة النظام 2,02، وجودة المعلومات 2,14، وجودة الواجهة 2,05، وهي قيم تقع جميعها دون الحد الأدنى لمعيار مقياس PSSUQ، مما يدل على مستوى استخدام ممتاز. أثبت النظام قدرته على تقديم توصيات ملائمة وفق تفضيلات المستخدم، وإن كان أداؤه يتأثر بمدى تجانس فئات الوجهات وطول الكلمات المفتاحية المدخلة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan sektor pariwisata menjadi salah satu indikator yang turut mendorong pertumbuhan ekonomi di berbagai negara, termasuk Indonesia. Selain berkontribusi terhadap perolehan devisa, sektor ini juga berperan dalam meningkatkan pendapatan daerah. Berdasarkan data Kementerian Pariwisata Republik Indonesia (2025), tercatat pada Juli 2025 jumlah wisatawan asing mengalami peningkatan mencapai 1.481.346 kunjungan, kemudian meningkat 13,01% dibandingkan Juli 2024. Sementara itu, perjalanan wisatawan domestik pada periode yang sama mencapai 101.200.000 perjalanan, dengan pertumbuhan 29,72% dibandingkan Juli 2024. Data ini menunjukkan bahwa sektor pariwisata di Indonesia terus mengalami pertumbuhan (Kementerian Pariwisata, 2025).

Peningkatan jumlah wisatawan tersebut menandakan tingginya minat masyarakat dalam melakukan perjalanan. Eddyono (2021) mendefinisikan pariwisata sebagai kegiatan perjalanan ke luar lingkungan tempat tinggal untuk rekreasi, bukan untuk bekerja atau mencari penghasilan. Berdasarkan fenomena ini, keberadaan sistem pendukung sangat dibutuhkan untuk memudahkan wisatawan memilih destinasi yang efisien dan selaras dengan kriteria preferensi personal pengguna.

Sistem rekomendasi dinilai sebagai salah satu pendekatan yang efektif guna menghasilkan usulan objek wisata yang selaras dengan kecenderungan minat pengguna. Menurut Jannach et al. (2011), *Content-Based Filtering* bekerja dengan

menghitung nilai *similarity* antara karakteristik *item* dan preferensi pengguna. Mekanisme kerja sistem ini bertumpu pada identifikasi fitur-fitur dari item yang telah diminati pengguna di masa lalu, yang kemudian dijadikan acuan untuk merekomendasikan item baru berkarakteristik sejenis. Profil pengguna dapat dibangun secara otomatis berdasarkan riwayat interaksi atau diperoleh secara eksplisit dari *input* pengguna, sedangkan deskripsi *item* mencakup atribut yang dapat diidentifikasi secara jelas (Astuti & Andriyani, 2025).

Metode *Content-Based Filtering* dipilih karena metode ini dinilai lebih cocok untuk konteks penelitian yang menggunakan data deskriptif serta tetap efektif meskipun data interaksi pengguna relatif terbatas. Ditinjau dari prasyarat datanya, algoritma *Collaborative Filtering* menunjukkan disparitas operasional terhadap *Content-Based Filtering*, dimana metode ini menuntut ketersediaan data entri nilai rating dari banyak pengguna (Pradana et al., 2022).

Beberapa penelitian telah mendukung keunggulan metode ini. Salsabilla et al. (2025) menggunakan pendekatan *Content-Based Filtering* yang dikombinasikan dengan TF-IDF dan *Cosine Similarity* untuk menyusun sistem rekomendasi wisata di Pekalongan. Hasilnya, sistem ini sukses menyajikan pilihan destinasi yang relevan dengan keinginan pengguna. Sementara itu, Yulfihani et al. (2024) menunjukkan penerapan *Content-Based Filtering* yang mempertimbangkan karakteristik objek wisata serta preferensi wisatawan dan mampu menghasilkan rekomendasi yang jauh lebih akurat.

Namun, sebagian besar riset terkait *Content-Based Filtering* untuk rekomendasi wisata ternyata masih membatasi diri dengan hanya mengandalkan

data konten dari objek wisata itu sendiri. Padahal, ulasan dan pengalaman nyata yang dibagikan wisatawan di media social menyimpan informasi yang kaya mengenai karakteristik suatu destinasi. Pemanfaatan ulasan pengguna untuk memperkaya konten berpotensi menghasilkan representasi destinasi yang lebih lengkap serta mencerminkan perspektif wisatawan yang sesungguhnya. Atas dasar itulah, penelitian ini mengembangkan dua fitur utama yaitu fitur rekomendasi berbasis *Content-Based Filtering* yang menghitung kemiripan antar destinasi menggunakan *Cosine Similarity*, serta fitur pencarian yang mengintegrasikan ulasan wisatawan ke dalam proses perhitungan TF-IDF dan *Cosine Similarity*, sehingga hasil pencarian tidak hanya bergantung pada konten destinasi melainkan juga dari pengalaman wisatawan yang pernah berkunjung.

Selain memberikan kontribusi pada bidang sistem rekomendasi dan kecerdasan buatan, penelitian ini juga mengintegrasikan nilai-nilai agama yang diharapkan mendorong manusia dalam merenungi keagungan dan keindahan ciptaan Allah SWT. Sebagaimana dalam QS. Ali Imran ayat 190–191:

﴿إِنَّ فِي خَلْقِ السَّمُوتِ وَالْأَرْضِ وَاخْتِلَافِ اللَّيْلِ وَالنَّهَارِ لَآيَاتٍ لِّأُولِي الْأَلْبَابِ ۚ ١٩٠ الَّذِينَ يَذْكُرُونَ اللَّهَ قِيَامًا وَقُعُودًا وَعَلَىٰ جُنُوبِهِمْ وَيَتَفَكَّرُونَ فِي خَلْقِ السَّمُوتِ وَالْأَرْضِ رَبَّنَا مَا خَلَقْتَ هَذَا بَاطِلًا سُبْحَانَكَ فَقِنَا عَذَابَ النَّارِ ١٩١﴾

“Sesungguhnya dalam penciptaan langit dan bumi, dan pergantian malam dan siang terdapat tanda-tanda bagi orang-orang yang berakal, (yaitu) orang-orang yang mengingat Allah sambil berdiri, duduk, dan berbaring, dan mereka memikirkan tentang penciptaan langit dan bumi seraya berkata, 'Ya Tuhan kami, tidaklah Engkau menciptakan ini dengan sia-sia; Maha Suci Engkau, maka peliharalah kami dari siksa neraka.'” (QS. Ali Imran: 190–191)

Mekanisme terciptanya alam semesta sekaligus sirkulasi pergantian siang dan malam merupakan tanda kekuasaan Allah yang nyata. Hamka (1990), dalam

Tafsir Al-Azhar memaparkan bahwa manusia yang menggunakan akalny akan membawa pengamatan alam ini ke tingkat refleksi jiwa. Hal ini membuat manusia sadar bahwa semua ciptaan Allah di bumi ini pasti memiliki maksud dan tujuannya masing-masing.

Penelitian ini memanfaatkan ulasan pengguna guna memahami preferensi wisatawan dapat dipandang sebagai bentuk refleksi terhadap keteraturan ciptaan Allah SWT. Melalui pendekatan ini, penelitian yang dilakukan pada akhirnya ikut berkontribusi dalam menyatukan pemikiran rasional dan kedalaman spiritual. Mengenai hal ini, Allah SWT menegaskan maknanya dalam QS. Al-Mulk ayat 15:

﴿هُوَ الَّذِي جَعَلَ لَكُمُ الْأَرْضَ ذَلُولًا فَامْشُوا فِي مَنَاكِبِهَا وَكُلُوا مِن رِّزْقِهِ ۚ وَإِلَيْهِ النُّشُورُ ۝ ١٥﴾

“Dialah yang menjadikan bumi itu mudah bagi kamu, maka berjalanlah di segala penjurunya dan makanlah sebagian dari rezeki-Nya. Dan hanya kepada-Nya-lah kamu (kembali setelah) dibangkitkan.” (QS. Al-Mulk: 15)

Melalui ayat tersebut, Allah SWT ingin menyampaikan pesan kepada manusia untuk menjadikan dunia mudah bagi manusia untuk ditempati dan dimanfaatkan, serta mengarahkan umat manusia agar bepergian ke segala penjuru dunia dalam rangka mencari rezeki. Dalam *Tafsir Ibnu Katsir* (2021b), dijelaskan bahwa perintah “berjalanlah di segala penjurunya” mengandung makna anjuran untuk berusaha, bekerja, dan mencari penghidupan, seperti berdagang dan melakukan aktivitas ekonomi lainnya, dengan tetap disertai sikap tawakal kepada Allah SWT.

Dengan demikian, ayat ini menunjukkan bahwa aktivitas eksplorasi dan pemanfaatan potensi alam, termasuk dalam bidang pariwisata, dapat dianggap

sebagai bentuk dari usaha manusia untuk mencari rezeki yang menjadi pemberian Allah SWT, selama dilakukan secara bijak dan bertanggung jawab. Lebih lanjut, keberagaman ini ditegaskan dalam QS. Al-Hujurat ayat 13:

﴿يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَىٰ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا ۚ إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَتْقَىٰكُمْ ۚ إِنَّ اللَّهَ عَلِيمٌ

حَبِيزٌ ١٣ ﴿﴾

“Wahai manusia, sesungguhnya Kami telah menciptakan kamu dari seorang laki-laki dan perempuan. Kemudian, Kami menjadikan kamu berbangsa-bangsa dan bersuku-suku agar kamu saling mengenal. Sesungguhnya yang paling mulia di antara kamu di sisi Allah adalah orang yang paling bertakwa. Sesungguhnya Allah Maha Mengetahui lagi Mahateliti.” (QS. Al-Hujurat: 13)

Melalui ayat ini, Kita diajarkan bahwa seluruh manusia lahir dari akar yang sama, yaitu dari Nabi Adam dan Hawa. Oleh karena itu, pada dasarnya setiap orang memiliki kedudukan yang sama. Tujuan Allah SWT menciptakan manusia berbeda bangsa dan suku adalah supaya mereka saling mengenal (*ta'aruf*), bukan untuk membedakan atau merendahkan satu sama lain. Dalam *Tafsir Ibnu Katsir* (2021a), dijelaskan bahwa perbedaan tersebut berfungsi sebagai sarana identifikasi dan interaksi sosial, sedangkan kemuliaan manusia di sisi Allah ditentukan oleh ketakwaannya.

Dalam konteks penelitian ini, perbedaan preferensi wisata antar pengguna merupakan bentuk keberagaman manusia. Maka dari itu, dibutuhkan sistem rekomendasi yang bisa memahami perbedaan secara adil. Pendekatan *Content-Based Filtering* digunakan dalam mengenali karakteristik pengguna, sehingga dapat menyajikan saran tempat wisata yang pas dan tepat untuk masing-masing pengguna. Selain itu, Rasulullah SAW juga mengingatkan dalam sabdanya:

مَنْ سَلَكَ طَرِيقًا يَلْتَمِسُ فِيهِ عِلْمًا سَهَّلَ اللَّهُ لَهُ بِهِ طَرِيقًا إِلَى الْجَنَّةِ

“Barangsiapa menempuh jalan untuk mencari ilmu, maka Allah akan memudahkan baginya jalan menuju surga.” (HR. Muslim no. 2699)

Hadist ini memperkuat bahwa menuntut ilmu dan melakukan penelitian yang bermanfaat bagi masyarakat merupakan bagian dari amal ibadah (2025). Dengan demikian, pengembangan sistem rekomendasi tujuan wisata ini diharapkan bisa membawa manfaat besar, baik dalam keilmuan secara akademik sekaligus memperkuat nilai-nilai spiritual.

1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang masalah di atas, maka permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana mengembangkan sistem rekomendasi destinasi wisata berbasis *Content-Based Filtering* menggunakan TF-IDF dan *Cosine Similarity*?
2. Bagaimana mengembangkan fitur pencarian destinasi wisata yang mengintegrasikan ulasan wisatawan dalam memberkaya konten?
3. Bagaimana kinerja sistem rekomendasi dalam menghasilkan rekomendasi destinasi wisata yang relevan berdasarkan *input* preferensi pengguna, serta bagaimana tingkat usability sistem dari sisi pengguna?

1.3 Batasan Masalah

Mengingat luasnya ruang lingkup dalam pengembangan sistem, penelitian ini berfokus pada batasan masalah yang dirumuskan di bawah ini:

1. Data ulasan hanya berupa teks dari *platform* media sosial X, TikTok dan YouTube, tanpa melibatkan data *non-teks* seperti gambar, video, atau *rating* bintang.
2. Ruang lingkup penelitian difokuskan pada 27 destinasi wisata di wilayah Jawa Timur sesuai dengan data yang diperoleh dari *website* resmi destinasi wisata seperti Sistem Informasi Daya Tarik Wisata Jawa Timur dan direktori pariwisata resmi sejenisnya.
3. Data ulasan pengguna digunakan untuk memperkaya konten dalam proses proses pencocokan kata kunci pada fitur pencarian, bukan sebagai data pelatihan model klasifikasi.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem rekomendasi destinasi wisata berbasis *Content-Based Filtering* menggunakan dengan TF-IDF dan *Cosine Similarity*.
2. Mengembangkan fitur pencarian destinasi yang mengintegrasikan data ulasan wisatawan untuk memperkaya konten
3. Mengevaluasi kinerja sistem rekomendasi dalam menghasilkan rekomendasi destinasi wisata yang relevan terhadap preferensi pengguna, serta mengevaluasi tingkat usability sistem dari sisi pengguna.

1.5 Manfaat Penelitian

Penelitian ini diharapkan mampu memberikan kontribusi dan manfaat nyata sebagai berikut:

1. Memberikan sistem rekomendasi destinasi wisata berbasis *Content-Based Filtering* yang mampu menghasilkan rekomendasi sesuai preferensi pengguna berdasarkan deskripsi destinasi.
2. Menyediakan informasi ulasan wisatawan sebagai gambaran persepsi pengguna terhadap kualitas suatu destinasi wisata.
3. Menjadi referensi akademik dalam pengembangan sistem rekomendasi wisata yang memuat fitur rekomendasi berbasis *Content-Based Filtering* dan fitur pencarian berbasis ulasan pengguna pada domain pariwisata.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terdahulu

Pengembangan sistem rekomendasi untuk tujuan wisata sebenarnya bukan hal baru karena sejumlah penelitian terdahulu sudah pernah membahasnya dengan berbagai metode dan pendekatan berbeda. Pendekatan yang cukup banyak digunakan adalah pendekatan *Content-Based Filtering* untuk memberikan saran wisata berdasarkan karakteristik objek wisata. Sebagian besar penelitian tersebut masih mengandalkan deskripsi destinasi dan kategori sebagai sumber konten tanpa memanfaatkan ulasan wisatawan yang sesungguhnya kaya akan informasi karakteristik destinasi.

Melalui penelitian berjudul “Pekalongan Regency Tourism Recommendation System with Content-Based Filtering”, Salsabilla dan Utomo (2025) mengembangkan sistem untuk merekomendasikan tempat wisata di Kabupaten Pekalongan. pendekatan yang dipakai merupakan *Content-Based Filtering*, lengkap beserta pembobotan TF-IDF serta kalkulasi *Cosine Similarity*. Sistem ini memproses delapan kategori wisata dari dataset resmi Dinas Pariwisata Pekalongan guna menilai kemiripan antar destinasi. Meskipun hasil riset membuktikan sistem ini bisa memberi rekomendasi yang relevan bagi pengguna, sayangnya pengembangannya masih terbatas pada kecocokan konten destinasi dan belum mempertimbangkan ulasan atau persepsi dari wisatawan.

Selanjutnya, Ilhamsyah et al. (2024) dalam penelitiannya yang berjudul “Pemodelan Sistem Rekomendasi Restoran Berdasarkan Preferensi Pengguna

dengan Pendekatan *Content-Based Filtering*” mengembangkan sistem rekomendasi restoran berdasarkan preferensi pengguna seperti rasa, harga, dan rating. Dataset dalam riset ini mencakup 661 ulasan restoran yang didapatkan melalui proses *web scraping* Google Maps via Apify. Dalam menghitung nilai *similarity* tiap item destinasi wisata, data yang terkumpul dianalisis menggunakan metode TF-IDF dan *Cosine Similarity*. Hasilnya, sistem mampu memunculkan pilihan wisata dengan rata-rata tingkat akurasi sebesar 73,33%.

Melalui penelitian berbeda, Pratiwi et al. (2025) membuat sebuah sistem rekomendasi untuk tempat wisata di Aceh Tamiang. Metode yang mereka gunakan adalah *Content-Based Filtering* dengan mengandalkan data ulasan objek wisata yang ada di Google Maps. Data dikumpulkan menggunakan SerpApi, kemudian melalui tahapan *Text Preprocessing*, dilanjutkan dengan pembobotan TF-IDF, dan diakhiri perhitungan *Cosine Similarity*. Sistem yang dihasilkan terbukti dapat memaparkan destinasi wisata yang relevan sesuai preferensi pengguna, khususnya pada atribut seperti "pantai", "sungai", dan "sepi".

Pada penelitian selanjutnya, Suryadi et al. (2025) mengembangkan penelitiang dengan judul “Sistem Rekomendasi Wisata Alam di Gunung Kidul Menggunakan Metode *Content-Based Filtering* Berbasis Web”, dengan menggabungkan informasi berupa kategori wisata, nilai *rating*, serta ulasan wisatawan sebagai dasar pembentukan rekomendasi. Hasil pengujian menunjukkan bahwa sistem tidak sekedar memudahkan pengguna dalam menentukan tempat liburan yang cocok, melainkan juga turut meningkatkan visibilitas destinasi-destinasi lokal.

Sementara itu, Maulana et al. (2026) mengembangkan sistem rekomendasi destinasi wisata di Provinsi Aceh dengan memanfaatkan algoritma *Cosine Similarity* serta *Haversine Formula*. Lewat cara ini, saran wisata yang dipaparkan tidak hanya berbasis pada kemiripan konten, namun memperhatikan jarak geografis antara pengguna dan lokasi wisata. Dataset yang dipakai mencakup 119 destinasi wisata dengan atribut nama, deskripsi, serta koordinat lokasi. Pendekatan ini pun terbukti sukses memberikan rekomendasi destinasi yang tepat, baik secara karakteristik konten maupun jarak fisik dari pengguna.

Di sisi lain Palelleng et al. (2025) membangun sistem edukasi wisata dengan konsep *Smart Tourism* untuk destinasi wisata di Toraja Utara, dengan menerapkan pendekatan *Content-Based Filtering* yang memanfaatkan TF-IDF dan *Cosine Similarity* dalam menghitung nilai *similarity* antar objek wisata. Melalui pengujian *User Acceptance Testing* (UAT), diperoleh angka kelayakan mencapai 89,66%. Hasil ini membuktikan bahwa aplikasi tersebut sudah sangat pantas untuk dimanfaatkan oleh para pelancong maupun pihak pengelola tempat wisata.

Selain pendekatan *Content-Based Filtering*, metode *Collaborative Filtering* sering diimplementasikan pada sistem rekomendasi wisata. Muahidin dan Azkiyah (2024) dalam penelitiannya yang berjudul "Penerapan Algoritma *Collaborative Filtering* pada Sistem Informasi Penyewaan Jasa *Tour Guide* Berbasis Web di Wisata Benang Kelambu" mengembangkan sistem informasi penyewaan jasa pemandu wisata menggunakan pendekatan *Waterfall* serta bahasa pemrograman PHP dan *framework* Laravel. Sistem rekomendasi pada penelitian ini dibangun berdasarkan *rating* dan ulasan dari pengguna lain untuk menampilkan *guide* yang

relevan dengan *guide* yang sedang dilihat atau disewa oleh pengguna. Berdasarkan uji coba lewat metode *black-box testing*, sistem terbukti mampu menjalankan semua fungsi utamanya dengan lancar dan bebas dari *error*. Namun, sistem ini mensyaratkan pengguna untuk melakukan registrasi dan login terlebih dahulu sebelum dapat melihat rekomendasi maupun memberikan ulasan terhadap *guide*.

Tabel 2.1 Penelitian Terdahulu

No	Peneliti	Judul	Metode	Objek Penelitian	Hasil
1	Salsabilla et al. (2025)	Sistem Rekomendasi Wisata Kabupaten Pekalongan dengan <i>Content-Based Filtering</i>	<i>Content-Based Filtering</i>	Destinasi wisata Kabupaten Pekalongan	Sistem berhasil memberikan rekomendasi destinasi wisata berdasarkan kesamaan kategori wisata serta memudahkan wisatawan dalam mencari tempat liburan yang sesuai dengan keinginan masing-masing..
2	Ilhamsyah et al. (2024)	Pemodelan Sistem Rekomendasi Restoran berdasarkan Preferensi Pengguna dengan Pendekatan <i>Content-Based Filtering</i>	<i>Content-Based Filtering</i>	Restoran	Nilai akurasi rata-rata sebesar 73,33% berhasil dicapai oleh sistem dalam menyajikan saran tempat makan yang pas dengan selera pengguna, mengacu pada data ulasan serta harga.
3	Pratiwi et al. (2025)	Penerapan Metode <i>Content-Based Filtering</i> dalam Sistem Rekomendasi Objek Wisata di Aceh Tamiang	<i>Content-Based Filtering</i>	Objek wisata di Aceh Tamiang (ulasan Google Maps)	Sistem dapat menyajikan rekomendasi yang relevan dengan preferensi pengguna berdasarkan analisis ulasan wisata seperti atribut pantai, sungai, dan suasana tempat.
4	Suryadi et al. (2025)	Sistem Rekomendasi Tempat Wisata Alam di Gunung Kidul Menggunakan Algoritma <i>Content-Based Filtering</i> Berbasis Web	<i>Content-Based Filtering</i>	Destinasi wisata alam di Gunung Kidul	Sistem dapat menyajikan saran tempat wisata yang dipersonalisasi berdasarkan kategori, deskripsi, rating, dan ulasan sehingga memudahkan wisatawan dalam menentukan tujuan wisata dan

No	Peneliti	Judul	Metode	Objek Penelitian	Hasil
					meningkatkan eksposur destinasi wisata lokal.
5	Maulana et al. (2026)	Sistem Rekomendasi Destinasi Tempat Wisata Provinsi Aceh menggunakan Metode <i>Content-Based Filtering</i>	<i>Content-Based Filtering</i>	Destinasi wisata di Provinsi Aceh	Sistem rekomendasi dapat menyajikan saran destinasi wisata yang mempertimbangkan kesamaan konten objek wisata serta kedekatan jarak geografis antara pengguna dan destinasi wisata.
6	Palelleng et al. (2025)	Pengembangan Aplikasi Edukasi Wisata Berbasis Smart Tourism dengan Sistem Rekomendasi <i>Content-Based Filtering</i>	<i>Content-Based Filtering</i>	Destinasi wisata di Toraja Utara	Sistem rekomendasi yang dikembangkan memperoleh nilai kelayakan sebesar 89,66% berdasarkan pengujian <i>User Acceptance Testing (UAT)</i> sehingga sudah sangat pantas dimanfaatkan oleh para wisatawan maupun pengelola tempat wisata.
7	Muahidin et al (2024)	Pengembangan Aplikasi Edukasi Wisata Berbasis Smart Tourism dengan Sistem Rekomendasi <i>Content-Based Filtering</i>	<i>Content-Based Filtering</i>	Destinasi wisata di Toraja Utara	Sistem rekomendasi tour guide berbasis <i>Collaborative Filtering</i> berhasil dibangun dan lolos uji <i>black-box</i> tanpa <i>error</i> . Kelemahan sistem ini yaitu fitur rekomendasi dan ulasan hanya dapat diakses setelah pengguna registrasi dan login.
8	Penelitian saat ini	Sistem Rekomendasi Tujuan Wisata di Jawa Timur Berbasis <i>Content-Based Filtering</i> Menggunakan <i>Cosine Similarity</i>	<i>Content-Based Filtering</i>	Destinasi wisata di Jawa Timur dan ulasannya	

Berdasarkan kajian terhadap penelitian terdahulu, sebagian besar penelitian sistem rekomendasi wisata masih berfokus pada pemanfaatan metode *Content-Based Filtering* dengan memanfaatkan deskripsi destinasi sebagai sumber konten

utama. Penelitian yang mengintegrasikan ulasan wisatawan untuk memperkaya konten dalam proses pencarian destinasi masih belum ada. Sehingga, penelitian ini juga mengembangkan fitur pencarian yang memanfaatkan ulasan wisatawan. Dengan ini, hasil pencarian tidak sekadar berpatokan pada data destinasi, melainkan turut mencerminkan pengalaman wisatawan.

2.2 Sistem Rekomendasi

Perkembangan teknologi menyebabkan jumlah informasi di internet meningkat secara signifikan. Kondisi ini sering menimbulkan *information overload*, yaitu keadaan ketika pengguna kesulitan menentukan pilihan karena banyaknya alternatif yang tersedia. Dalam membantu pengguna dalam mengambil keputusan, dikembangkan sistem rekomendasi yang dapat menyajikan pilihan objek wisata secara cepat, tepat, dan relevan.

Menurut Novianti et al. (2021), sistem rekomendasi merupakan sebuah alat yang berfungsi menyajikan pilihan atau saran *item* kepada pengguna dengan mengacu pada profil pengguna maupun karakteristik *item* yang tersedia. Dalam konteks pariwisata, sistem mempermudah wisatawan dalam menentukan tempat liburan yang sesuai dengan selera mereka tanpa harus menelusuri seluruh informasi destinasi yang tersedia.

Jannach et al. (2011) menjelaskan bahwa pada dasarnya ada tiga metode utama yang sering digunakan dalam membangun sebuah sistem rekomendasi, yakni *Content-Based Filtering*, *Collaborative Filtering*, serta *Hybrid Recommendation*. *Content-Based Filtering* bekerja dengan menghasilkan rekomendasi dengan membandingkan kesamaan antara atribut *item* dan preferensi pengguna. Sebagai

contoh, pengguna yang memiliki minat pada wisata alam akan lebih banyak direkomendasikan destinasi seperti gunung atau pantai (Yaya Suharya et al., 2021).

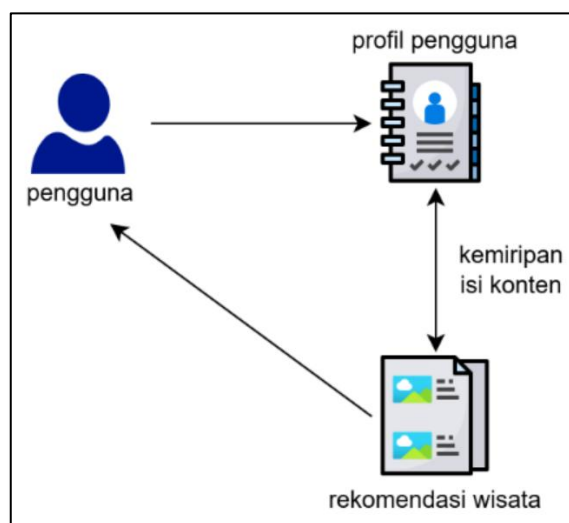
Pendekatan *Collaborative Filtering* bekerja dengan cara melihat kesamaan perilaku antar pengguna. Melalui metode ini, saran yang muncul didasarkan kepada seleda pengguna lain yang memiliki ketertarikan yang serupa (Howay & Rianto, 2021). Sementara itu, *Hybrid Recommendation* hadir sebagai tengkin yang memadukan dua pendekatan atau lebih. Tujuannya adalah untuk saling melengkapi kekurangan dari tiap metode sehingga mampu menghasilkan rekomendasi yang jauh lebih akurat (Putri et al., 2021).

Metode *Content-Based Filtering* diterapkan dalam proses perancangan sistem ini, karena metode tersebut sesuai dengan karakteristik data yang digunakan, yaitu deskripsi dan ulasan destinasi wisata.

2.3 *Content-Based Filtering*

Content-Based Filtering dikenal sebagai pendekatan yang menghasilkan rekomendasi dengan cara menganalisis kesamaan konten antar *item*. Prinsip utama metode ini adalah “jika pengguna menyukai suatu *item*, maka ia kemungkinan besar akan menyukai *item* lain yang serupa” (2019). Dalam penerapannya, metode *Content-Based Filtering* melibatkan dua elemen utama, yakni data profil pengguna serta karakteristik *item* yang akan direkomendasikan. *Item* direpresentasikan melalui atribut atau fitur tertentu yang menggambarkan karakteristik dari *item* tersebut, sedangkan profil pengguna menggambarkan preferensi atau minat pengguna terhadap jenis *item* tertentu.

Pada penelitian ini, *item* yang digunakan berupa destinasi wisata di Jawa Timur yang direpresentasikan melalui atribut kategori wisata, lokasi, serta deskripsi destinasi. Sementara itu, profil pengguna direpresentasikan melalui preferensi yang *diinputkan* oleh pengguna dalam sistem. Berdasarkan kedua komponen tersebut, sistem kemudian mengukur nilai kemiripan antara profil pengguna dengan konten destinasi wisata yang kemudian menghasilkan rekomendasi destinasi yang paling relevan. Skema mekanisme rekomendasi menggunakan metode *Content-Based Filtering* disajikan dalam Gambar 2.1.



Gambar 2.1 Mekanisme *Content-Based Filtering*

Berdasarkan Gambar 2.1, proses dimulai dari pembentukan profil pengguna atau riwayat aktivitas pengguna yang merepresentasikan preferensi pengguna. Selanjutnya sistem mengukur kemiripan antara profil pengguna dengan konten wisata. Destinasi dengan nilai *similarity* tertinggi kemudian disajikan kepada pengguna sebagai rekomendasi opsi wisata. Dalam menghitung nilai *similarity* antara profil pengguna dan destinasi wisata, penelitian ini menerapkan representasi

dokumen berbasis pembobotan kata dengan metode TF-IDF dan perhitungan *similarity* dengan metode *Cosine Similarity*.

Pendekatan *Content-Based Filtering* memiliki keunggulan karena dapat menghasilkan rekomendasi meskipun riwayat aktivitas atau penilaian pengguna lain terbatas. Meski begitu, keterbatasan dari pendekatan ini ialah potensi *over-specialization*, yaitu kondisi ketika sistem hanya memberikan saran *item* yang sangat mirip dengan *item* yang sebelumnya dipilih oleh pengguna (Oyadila et al., 2025). Selain fitur rekomendasi, penelitian ini juga mengembangkan fitur pencarian yang memanfaatkan ulasan wisatawan sebagai pengayaan konten, sehingga representasi destinasi menjadi lebih lengkap dan mencerminkan pengalaman wisatawan.

2.4 *Text Preprocessing*

Text Preprocessing menjadi langkah pertama dalam pengolahan data untuk membersihkan serta menstandarkan teks sebelum melalui proses analisis lebih lanjut. Tahapan ini diterapkan dalam penelitian baik pada fitur rekomendasi maupun fitur pencarian. Beberapa langkah utama yang digunakan pada tahap ini meliputi:

1. *Cleaning*

Tahap ini melalui pembersihan komponen tidak relevan seperti karakter khusus, *emoticon*, angka, tanda baca, spasi berlebih, tautan, serta tagar yang muncul di dalam teks. Tujuannya adalah untuk menghilangkan *noise* yang dapat mengganggu proses analisis selanjutnya, sehingga data menjadi lebih murni.

2. *Case folding*

Dalam langkah ini, dilakukan dengan mengonversi semua huruf dalam seluruh dokumen menjadi bentuk *lowercase*, supaya menjaga konsistensi penulisan kata dan menghindari duplikasi token akibat perbedaan kapitalisasi, misalnya "Pantai" dan "pantai" dianggap sama.

3. *Tokenizing*

Langkah *tokenizing* ini memecah kalimat atau paragraf panjang menjadi token, yang menjadi fondasi untuk analisis lebih lanjut. Proses ini mempermudah analisis teks seperti perhitungan frekuensi kata.

4. *Normalization*

Normalization bertujuan untuk mengganti kata yang mengandung kesalahan penulisan, *slang*, bahasa daerah, atau variasi *informal* dengan bentuk standar berdasarkan Kamus Besar Bahasa Indonesia (KBBI).

5. *Stopword removal*

Langkah ini dilakukan dalam rangka menyaring kata umum atau *stopwords* yang tidak memiliki makna penting dalam analisis teks, seperti kata penghubung ("dan", "atau"), kata ganti ("saya", "ini"), atau kata filler lainnya. Proses ini memfokuskan analisis pada kata yang benar-benar bermakna dalam konteks sentimen.

6. *Stemming*

Langkah ini mengubah kata yang berimbuhan menjadi bentuk dasar atau akar kata.

2.5 *Term Frequency–Inverse Document Frequency (TF-IDF)*

Dalam pemrosesan teks, pendekatan TF-IDF digunakan dalam menentukan bobot prioritas suatu *term* pada dokumen melalui representasi matriks numerik. Metode ini memegang peranan krusial dalam arsitektur *text mining* serta *information retrieval*. Secara matematis, kalkulasi TF-IDF ditopang oleh variabel TF dan IDF. Komponen TF merepresentasikan rasio kemunculan suatu kata terhadap total kata pada dokumen bersangkutan. Di sisi lain, komponen IDF mengevaluasi distribusi kata tersebut pada seluruh koleksi dokumen di dalam database; kata yang bersifat eksklusif (jarang muncul) akan dianugerahi nilai IDF yang tinggi karena faktor informatifnya, sedangkan kata yang tersebar luas secara umum akan ditekan nilai IDF-nya menjadi lebih rendah (Santoti et al., 2025).

Bobot TF-IDF didapatkan dari hasil perkalian nilai TF dan IDF yang dinyatakan oleh persamaan berikut:

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (2.1)$$

Adapun perhitungan TF dan IDF secara matematis masing-masing dinyatakan sebagai berikut:

$$TF(t, d) = \frac{f_{t,d}}{N_d} \quad (2.2)$$

$$IDF(t) = \log\left(\frac{N}{df_t}\right) \quad (2.3)$$

di mana:

- $f_{t,d}$: frekuensi kemunculan kata t dalam dokumen d
- N_d : jumlah kata dalam dokumen d
- N : jumlah total dokumen
- df_t : jumlah dokumen yang mengandung kata t

Melalui penelitian ini, TF-IDF digunakan untuk dua keperluan, yaitu merepresentasikan konten destinasi wisata dalam bentuk vektor pada fitur rekomendasi, serta menghitung bobot kata kunci dan konten destinasi secara bersamaan pada fitur pencarian sebelum dilakukan perhitungan *similarity* dengan *Cosine Similarity*.

2.6 *Cosine Similarity*

Setelah bobot TF-IDF dari setiap dokumen diperoleh, langkah selanjutnya adalah menghitung tingkat kemiripan antar dokumen. Pendekatan yang umum digunakan dalam menghitung *similarity* adalah *Cosine Similarity*. Pendekatan ini dipakai dalam menghitung nilai *similarity* antara dokumen dengan membandingkan vektor representasi dari tiap dokumen (Santoti et al., 2025).

Dalam menghitung nilai *Cosine Similarity* digunakan rumus persamaan berikut:

$$\text{Cosine Similarity}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} \quad (2.4)$$

Keterangan:

$A \cdot B$: hasil *dot product* antara dua vektor TF-IDF
 $\|A\|$ dan $\|B\|$: panjang (*norm*) dari masing-masing vektor

Nilai $A \cdot B$ merupakan hasil *dot product* antara dua vektor yang diperoleh dengan mengalikan setiap elemen yang bersesuaian dan menjumlahkannya, yang secara matematis dirumuskan sebagai berikut:

$$A \cdot B = \sum_{i=1}^n (A_i \times B_i) \quad (2.5)$$

Sementara itu, panjang vektor atau norma digunakan sebagai faktor normalisasi yang dihitung dari akar jumlah kuadrat setiap elemen vektor, sebagaimana dinyatakan pada persamaan berikut:

$$\|A\| = \sqrt{\sum_{i=1}^n (w_i^2)} \quad (2.6)$$

Nilai *Cosine Similarity* bernilai antara nol hingga satu, dengan nilai yang mendekati satu menandakan tingkat kemiripan yang tinggi. Di sisi lain, nilai yang mendekati nol menandakan tingkat kemiripan yang rendah terhadap dua dokumen. Dalam penelitian ini pendekatan *Cosine Similarity* diterapkan dalam dua keperluan. Pertama, pada fitur rekomendasi, *Cosine Similarity* dipakai dalam menghitung kemiripan antara vektor TF-IDF destinasi yang dipilih pengguna dengan vektor destinasi lainnya. Kedua, pada fitur pencarian, *Cosine Similarity* dipakai dalam menghitung kemiripan antara vektor kata kunci pengguna dengan vektor konten destinasi yang telah diperkaya dengan ulasan wisatawan.

2.7 Evaluasi Sistem

Untuk mengukur tingkat keberhasilan aplikasi yang telah selesai dibuat, pelaksanaan evaluasi menjadi langkah wajib yang harus ditempuh dalam penelitian ini. Proses evaluasi sistem mencakup dua aspek utama, yaitu evaluasi sistem rekomendasi dan evaluasi *usability* dari sisi pengguna. Evaluasi sistem rekomendasi sendiri digunakan untuk menilai relevansi dari hasil rekomendasi, sedangkan evaluasi *usability* diperlukan supaya mengetahui tingkat kepuasan pengguna terhadap sistem.

2.7.1 *Precision@K*

Evaluasi sistem rekomendasi dilakukan dengan memanfaatkan metrik *Precision@K*. *Precision@K* merupakan metode evaluasi untuk mengukur tingkat presisi hasil rekomendasi berdasarkan K item teratas yang disajikan oleh sistem. Metode ini hanya mempertimbangkan *item* yang berada pada posisi teratas hingga batas K, sedangkan *item* di luar peringkat tersebut tidak diperhitungkan. *Precision@K* digunakan untuk mengetahui seberapa banyak *item* relevan yang berhasil direkomendasikan oleh sistem pada hasil teratas. Relevansi hasil rekomendasi ditentukan berdasarkan penilaian *anotator* terhadap kesesuaian antara hasil rekomendasi dengan destinasi *input*.

$$Precision@K = \frac{\text{Jumlah destinasi relevan pada top} - K}{K} \quad (2.7)$$

Nilai *Precision@K* bernilai antara 0 hingga 1. Semakin tinggi nilai *Precision@K*, yang menandakan bahwa kualitas rekomendasi yang dihasilkan semakin baik karena menunjukkan bahwa semakin banyak *item* relevan yang muncul pada hasil rekomendasi teratas.

2.7.2 *Post-Study System Usability Questionnaire (PSSUQ)*

Evaluasi *usability* menggunakan metode PSSUQ. PSSUQ dikenal sebagai instrumen yang diterapkan dalam mengukur tingkat kepuasan pengguna terhadap suatu sistem setelah digunakan serta menilai aspek *usability* secara kuantitatif.

PSSUQ awalnya dikembangkan oleh IBM guna mendukung proyek *System Usability MetricS* (SUMS). Melalui serangkaian pembaruan, kuesioner ini berevolusi hingga mencapai versi ketiga (PSSUQ Version 3) yang terdiri dari 16

butir pertanyaan. Versi ini disusun untuk meningkatkan reliabilitas dan validitas pengukuran *usability* dengan mengeliminasi *item* yang kurang konsisten pada versi sebelumnya (Suwandy et al., 2022). PSSUQ mengukur *usability* berdasarkan tiga aspek utama, yaitu:

1. *System Usefulness* (SysUse)

Aspek ini mengukur kualitas sistem dari segi kemudahan penggunaan, efisiensi, dan kemampuan sistem dalam membantu pengguna menyelesaikan tugas. Nilai *System Usefulness* diperoleh dari rata-rata skor pertanyaan 1 sampai dengan 6.

2. *Information Quality* (InfoQual)

Aspek ini menilai kualitas informasi yang disajikan oleh sistem, meliputi kejelasan, kelengkapan, serta kemudahan dalam memahami informasi, dengan nilai *Information Quality* yang dihitung dari rata-rata skor pertanyaan 7 sampai 12.

3. *Interface Quality* (IntQual)

Aspek ini menilai tampilan desain dari sistem, termasuk kenyamanan visual, konsistensi desain, dan kemudahan navigasi. Nilai *Interface Quality* diperoleh dari rata-rata skor pertanyaan 13 sampai dengan 15.

Selain ketiga aspek tersebut, PSSUQ juga menghasilkan nilai *Overall*, yaitu nilai keseluruhan *usability* sistem yang dihitung berdasarkan rata-rata skor dari seluruh pertanyaan (1–16). Pengukuran dalam PSSUQ memakai skala *likert* dengan rentang nilai 1 sampai 7 untuk merepresentasikan tingkat persetujuan responden terhadap setiap pernyataan. Skor yang digunakan meliputi nilai 1 untuk Sangat Setuju (SS), 2 untuk Setuju (S), 3 untuk Cukup Setuju (CS), 4 untuk Netral (N), 5

untuk Cukup Tidak Setuju (CTS), 6 untuk Tidak Setuju (TS), dan 7 untuk Sangat Tidak Setuju (STS). Semakin rendah skor yang diperoleh, semakin tinggi tingkat kepuasan dan *usability* sistem menurut persepsi pengguna (Indriani et al., 2024). Perhitungan dilakukan menggunakan rata-rata:

$$\bar{X} = \frac{\sum x}{N} \quad (2.8)$$

Keterangan:

$\bar{X}$: Nilai rata-rata setiap aspek
 $\sum x$: Total skor penilai responden setiap sub-skala
 N : Jumlah *item* pertanyaan setiap sub-skala

Rumus tersebut digunakan untuk menghitung nilai rata-rata dengan menjumlahkan seluruh skor jawaban responden pada setiap sub-skala, kemudian membaginya dengan jumlah *item* pertanyaan. Nilai rata-rata yang diperoleh digunakan untuk menginterpretasikan tingkat kepuasan pengguna, di mana semakin rendah skor menunjukkan tingkat *usability* sistem yang semakin baik.

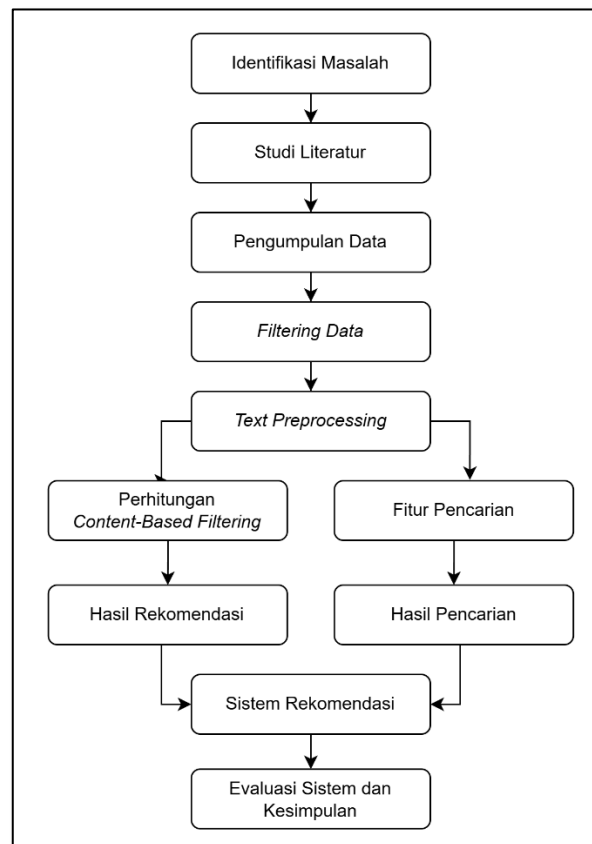
BAB III

DESAIN DAN IMPLEMENTASI

3.1 Desain Penelitian

Penelitian ini termasuk dalam kategori penelitian terapan (*applied research*) karena berfokus pada penerapan teori dan metode ilmiah dalam pengembangan sistem rekomendasi wisata berbasis teknologi. Hasil penelitian ini diharapkan mampu memberikan solusi praktis bagi wisatawan dalam menentukan destinasi wisata yang sesuai dengan preferensi masing-masing. Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen, karena melibatkan pengumpulan data empiris, penerapan metode *Content-Based Filtering*, serta evaluasi terhadap sistem rekomendasi yang dihasilkan.

Desain penelitian ini menjelaskan langkah-langkah yang dilakukan secara sistematis guna membangun sistem rekomendasi wisata berbasis *Content-Based Filtering*. Setiap tahapan dirancang secara sistematis mulai dari identifikasi masalah, pengumpulan dan pengolahan data, hingga evaluasi sistem dan penarikan kesimpulan, sehingga proses penelitian dapat berjalan secara terstruktur dan terarah untuk mencapai hasil akhir yang relevan. Gambar 3.1 berikut menunjukkan diagram alur penelitian yang digunakan dalam penelitian ini.



Gambar 3.1 Diagram Alur Penelitian

Merujuk pada Gambar 3.1, sistematika penelitian ini diawali dengan tahapan identifikasi masalah dan studi literatur untuk memahami permasalahan serta memperoleh landasan teori yang relevan dengan sistem rekomendasi. Selanjutnya dilakukan pengumpulan data yang mencakup data deskripsi destinasi wisata dan data ulasan pengguna dari beberapa *platform* media sosial. Data yang telah diperoleh kemudian disaring melalui proses *filtering* untuk memastikan bahwa hanya data yang relevan digunakan dalam penelitian.

Selanjutnya, data diproses melalui tahap *Text Preprocessing* untuk membersihkan dan menstandarkan teks agar dapat diolah secara komputasional. Data yang telah diproses kemudian digunakan dalam dua proses utama, yaitu

perhitungan *Content-Based Filtering* untuk menghasilkan rekomendasi destinasi wisata berdasarkan kesamaan konten, serta fitur pencarian yang mengintegrasikan ulasan wisatawan untuk memperkaya konten. Hasil dari kedua proses tersebut selanjutnya diintegrasikan ke dalam sistem rekomendasi, sehingga pengguna dapat memperoleh rekomendasi destinasi yang relevan dengan preferensi pribadi maupun hasil pencarian berdasarkan kata kunci yang *diinputkan*. Tahap akhir penelitian adalah evaluasi sistem dan penarikan kesimpulan.

3.2 Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan melalui metode *web scraping*, yaitu pengambilan data secara otomatis dari beberapa *platform* media sosial yang memuat opini pengguna terkait destinasi wisata di Jawa Timur. *Platform* yang digunakan meliputi X, TikTok, dan YouTube.

Pengambilan data dari *platform* X dilakukan menggunakan Tweet Harvest, yaitu *tool* berbasis Python yang memanfaatkan API publik untuk mengekstraksi tweet yang berisi ulasan atau opini pengguna. Untuk *platform* TikTok, data komentar diperoleh menggunakan layanan Apify melalui fitur TikTok *Comments Scraper* dengan memasukkan URL video yang relevan dengan destinasi wisata. Sementara itu, komentar dari YouTube dikumpulkan menggunakan *YouTube Data API V3* dari video yang berkaitan dengan topik wisata di Jawa Timur.

Data komentar yang telah diperoleh kemudian melalui tahap seleksi untuk memastikan kesesuaiannya dengan konteks ulasan destinasi wisata sebelum digunakan pada tahap selanjutnya. Selain itu, aspek etika penelitian juga diperhatikan dengan hanya menggunakan data yang bersifat publik serta tidak

menyertakan informasi pribadi pengguna. Data yang digunakan terbatas pada teks komentar atau ulasan tanpa menyimpan identitas akun, sehingga proses pengumpulan data tetap bertanggung jawab dan sesuai dengan prinsip etika penelitian.

3.2.1 Data Penelitian

Data yang digunakan terdiri dari dua jenis, yaitu *dataset* konten destinasi wisata dan *dataset* ulasan pengguna. *Dataset* konten digunakan dalam proses *Content-Based Filtering*, sedangkan *dataset* ulasan digunakan untuk memperkaya konten pada fitur pencarian destinasi wisata.

Dataset konten destinasi wisata diperoleh dari sumber sekunder seperti *website* resmi Sistem Informasi Daya Tarik Wisata Jawa Timur, direktori pariwisata, serta *website* resmi lain yang menyediakan informasi terkait destinasi wisata. Proses pengambilan data konten destinasi dilakukan pada tanggal 26 Januari 2026. Data yang dikumpulkan meliputi informasi nama destinasi, lokasi, kategori, dan deskripsi destinasi wisata di wilayah Jawa Timur.

Sementara itu, data ulasan pengguna diperoleh dari *platform* media sosial X, TikTok, dan YouTube, dalam rentang waktu 19 November 2025 hingga 23 Januari 2026, yang berisi opini pengguna terkait destinasi wisata di Jawa Timur. Data yang telah dikumpulkan kemudian melalui proses seleksi sebelum digunakan pada tahap selanjutnya.

3.2.1.1 Dataset Deskripsi Destinasi (Konten)

Dataset deskripsi destinasi digunakan untuk membentuk *profil item* (destinasi wisata) yang menjadi dasar dalam metode *Content-Based Filtering*. *Dataset* ini berisi 27 data destinasi wisata di Jawa Timur, dengan setiap destinasi memuat informasi dasar dan deskriptif mengenai destinasi tersebut, seperti nama destinasi, lokasi, kategori wisata dan deskripsi yang menggambarkan daya tarik destinasi. Data ini diperoleh melalui pengumpulan manual dari berbagai sumber daring terpercaya, seperti situs resmi Dinas Pariwisata Provinsi Jawa Timur, Google Maps, dan berbagai sumber yang berkaitan. Sampel data deskripsi konten disajikan pada Tabel 3.1.

Tabel 3.1 Sampel Data Deskripsi Destinasi Wisata

Fitur	Penjelasan	Contoh
Nama Destinasi	Nama objek destinasi wisata	Gunung Bromo
Lokasi	Lokasi destinasi wisata di wilayah Jawa Timur	Probolinggo
Kategori Wisata	Jenis wisata	wisata alam pegunungan dan air terjun
Deskripsi	Uraian mengenai destinasi wisata	Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Destinasi ini dikenal dengan lanskap pegunungan vulkanik, lautan pasir luas, serta kawah aktif yang menjadi daya tarik utama wisata alam. Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo, Pasir Berbisik, dan Bukit Teletubbies.

Tabel 3.1 menunjukkan struktur *dataset* deskripsi destinasi wisata yang digunakan dalam penelitian. *Dataset* tersebut terdiri dari empat atribut utama, yaitu nama destinasi, lokasi, kategori wisata, dan deskripsi destinasi. Informasi deskripsi tersebut digunakan sebagai konten utama dalam proses *Content-Based Filtering* untuk membentuk *profil item* dan menghitung tingkat kemiripan antar destinasi wisata.

3.2.1.2 Dataset Ulasan Pengguna

Dataset ulasan pengguna berisi opini dan pengalaman wisatawan terhadap destinasi wisata di Jawa Timur, yang digunakan untuk memperkaya konten dalam fitur pencarian. Data dikumpulkan secara sekunder dari *platform* media sosial X, TikTok dan YouTube, karena *platform* tersebut menyediakan opini pengguna yang bersifat terbuka dan dapat diakses publik. Proses pengumpulan data dilakukan dengan cara menelusuri kata kunci berupa nama dari tiap destinasi wisata untuk memperoleh ulasan yang relevan. Setiap destinasi diambil sejumlah 25 ulasan pengguna, sehingga total data yang diperoleh adalah 675 ulasan, yang cukup mewakili persepsi wisatawan terhadap tiap destinasi.

Setelah data ulasan terkumpul, dilakukan proses pelabelan sentimen menggunakan metode *crowdsourcing*. Pada tahap ini, setiap ulasan dilabeli oleh lima orang penilai (*anotator*) untuk menentukan apakah isi ulasan tersebut bernada positif atau negatif. Penilai merupakan mahasiswa yang memiliki pemahaman terhadap konteks bahasa indonesia. Hasil akhir label ditetapkan berdasarkan pendapat mayoritas. Data ulasan yang telah berlabel ini digunakan untuk memperkaya konten dalam fitur pencarian, sehingga sistem dapat memanfaatkan informasi dari pengalaman nyata wisatawan dalam menghasilkan hasil pencarian yang lebih relevan. Contoh data ulasan pengguna dipaparkan pada Tabel 3.2 berikut.

Tabel 3.2 Contoh Sampel Data Ulasan

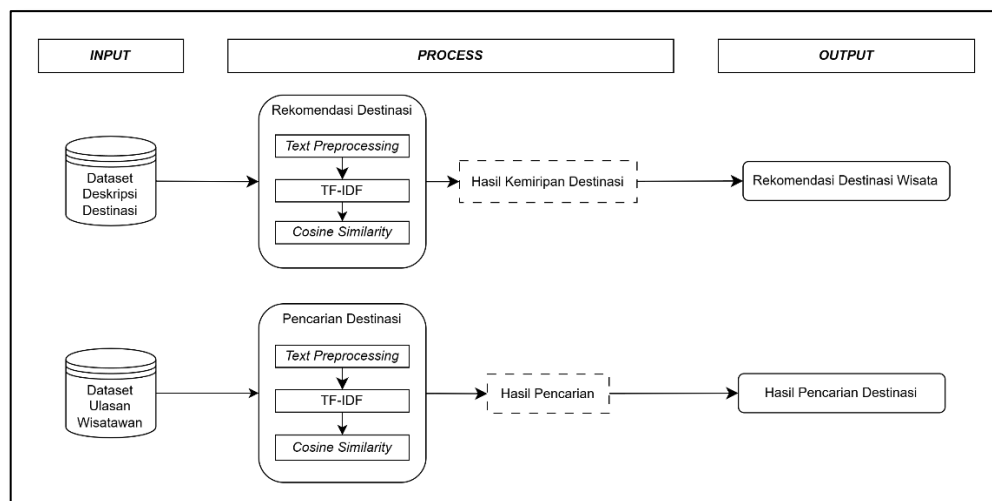
No	Nama Destinasi	Ulasan
1	Pantai Papuma	Papuma emang secantik itu 😍
2	Pantai Papuma	ke papuma butuh kemampuan dan kesabaran ekstra menghadapi kondisi jalan yang jelek

No	Nama Destinasi	Ulasan
3	Gunung Bromo	Pesona gunung Bromo memiliki daya tarik untuk wisatawan datang berkunjung.. Terutama sunrise nya
4	Gunung Bromo	pengalaman pertama ke Bromo,,,,, wow MasyaAllah seindah itu ciptaan Allah
5	Kawah Ijen	Jalanya udah mulus banget dulu ndaki ke kawah ijen tahun 2016 jalan utuh layaknya pegunungan
6	Kawah Ijen	Tanjakannya cukup tajam dan harus hati hati

Ulasan-ulasan di atas menunjukkan bahwa data yang digunakan bersifat *informal*, banyak mengandung kata tidak baku, singkatan, serta ekspresi emosional seperti *emoji*. Oleh karena itu, data perlu diproses melalui tahap *Text Preprocessing* supaya teks siap diproses pada tahap berikutnya.

3.3 Desain Sistem

Pada penelitian ini, Gambar 3.2 menyajikan mekanisme kerja sistem rekomendasi wisata yang dikembangkan, mulai dari *input* data hingga menghasilkan rekomendasi akhir kepada pengguna.



Gambar 3.2 Diagram Desain Sistem

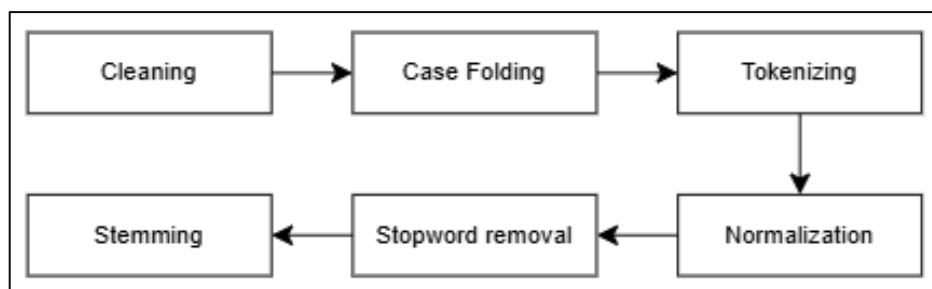
Berdasarkan Gambar 3.2, metode *Content-Based Filtering* digunakan sebagai metode utama dalam menghasilkan rekomendasi destinasi wisata

berdasarkan kesamaan konten destinasi. Bukan hanya itu, sistem menyediakan fitur pencarian yang memanfaatkan ulasan wisatawan untuk memperkaya konten, sehingga hasil pencarian tidak hanya bergantung pada konten destinasi, namun juga mencerminkan pengalaman dari wisatawan.

3.4 Tahap *Text Preprocessing*

Setelah data terkumpul, tahap selanjutnya adalah melakukan tahap *Text Preprocessing* agar siap digunakan dalam proses sistem rekomendasi. Tahap ini ditujukan untuk memfilter, menormalisasi, serta mempersiapkan data teks sehingga dapat diolah secara komputasional.

Sebelum dilakukan tahapan *Text Preprocessing*, dilakukan proses *filtering data* untuk memastikan bahwa hanya data yang relevan dan sesuai dengan konteks penelitian yang digunakan. Pada tahap ini, ulasan yang tidak berisi opini pengguna, komentar kosong, iklan, atau tautan promosi akan dihapus dari *dataset*. Proses *filtering* ini penting agar data yang digunakan benar-benar mencerminkan pengalaman wisatawan dan informasi destinasi secara *valid* sebelum masuk ke tahap *Text Preprocessing*.



Gambar 3.3 Diagram Tahapan *Text Preprocessing*

dalam penelitian ini dilakukan secara berurutan melalui beberapa langkah utama, yang terdiri atas *cleaning*, *case folding*, *tokenizing*, *normalization*, *stopword removal*, dan *stemming*. Rangkaian proses ini diterapkan guna mentransformasikan data teks tidak terstruktur menjadi representasi data yang lebih sistematis dan siap dianalisis pada tahap pembobotan. Adapun penjelasan mendalam mengenai masing-masing tahapan *text preprocessing* tersebut dijabarkan sebagai berikut:

1. *Cleaning*

Tahapan *cleaning* dilakukan untuk mengeliminasi elemen-elemen pengganggu (*noise*) yang tidak memiliki bobot informasi dalam pemrosesan teks. Proses ini meliputi penghapusan karakter khusus, simbol *emoticon*, angka, tanda baca, spasi berlebih, karakter baris baru, tautan URL, serta tanda *hashtag* yang terkandung di dalam data teks asli.

Tabel 3.3 Hasil *Text Preprocessing Cleaning*

Data Sebelum Cleaning	Data Sesudah Cleaning
probolinggo	probolinggo
wisata alam pegunungan dan air terjun	wisata alam pegunungan dan air terjun
Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Destinasi ini dikenal dengan lanskap pegunungan vulkanik, lautan pasir luas, serta kawah aktif yang menjadi daya tarik utama wisata alam. Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo, Pasir Berbisik, dan Bukit Teletubbies.	Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru Destinasi ini dikenal dengan lanskap pegunungan vulkanik lautan pasir luas serta kawah aktif yang menjadi daya tarik utama wisata alam Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo Pasir Berbisik dan Bukit Teletubbies

2. *Case folding*

Tahapan *case folding* dilakukan dengan mengonversi seluruh karakter huruf menjadi huruf kecil (*lowercase*). Proses ini bertujuan guna menormalisasi data teks dan menghindari terjadinya duplikasi atau perbedaan makna *term* yang disebabkan oleh variasi penggunaan huruf kapital.

Tabel 3.4 Hasil *Text Preprocessing Case Folding*

Data Sebelum Case Folding	Data Sesudah Case Folding
probolinggo	probolinggo
wisata alam pegunungan dan air terjun	wisata alam pegunungan dan air terjun
Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru Destinasi ini dikenal dengan lanskap pegunungan vulkanik lautan pasir luas serta kawah aktif yang menjadi daya tarik utama wisata alam Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo Pasir Berbisik dan Bukit Teletubbies	gunung bromo merupakan gunung berapi aktif di jawa timur yang berada dalam kawasan taman nasional bromo tengger semeru destinasi ini dikenal dengan lanskap pegunungan vulkanik lautan pasir luas serta kawah aktif yang menjadi daya tarik utama wisata alam beberapa lokasi populer di kawasan ini antara lain kawah bromo pasir berbisik dan bukit teletubbies

3. Tokenizing

Teks dipecah menjadi potongan-potongan kata (*token*) agar dapat diolah lebih lanjut.

Tabel 3.5 Hasil *Text Processing Tokenizing*

Data Sebelum Tokenizing	Data Sesudah Tokenizing
probolinggo	['probolinggo']
wisata alam pegunungan dan air terjun	['wisata', 'alam', 'pegunungan', 'dan', 'air', 'terjun']
gunung bromo merupakan gunung berapi aktif di jawa timur yang berada dalam kawasan taman nasional bromo tengger semeru destinasi ini dikenal dengan lanskap pegunungan vulkanik lautan pasir luas serta kawah aktif yang menjadi daya tarik utama wisata alam beberapa lokasi populer di kawasan ini antara lain kawah bromo pasir berbisik dan bukit teletubbies	['gunung', 'bromo', 'merupakan', 'gunung', 'berapi', 'aktif', 'di', 'jawa', 'timur', 'yang', 'berada', 'dalam', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'ini', 'dikenal', 'dengan', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'serta', 'kawah', 'aktif', 'yang', 'menjadi', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'beberapa', 'lokasi', 'populer', 'di', 'kawasan', 'ini', 'antara', 'lain', 'kawah', 'bromo', 'pasir', 'berbisik', 'dan', 'bukit', 'teletubbies']

4. Normalization

Proses transformasi kata tidak baku, istilah gaul (*slang*), maupun singkatan menjadi bentuk kosakata yang baku disebut sebagai *normalization*. Pada penelitian ini, penyesuaian kata tersebut merujuk pada ketentuan Kamus Besar Bahasa

Indonesia (KBBI) guna memastikan konsistensi seluruh dokumen teks sebelum memasuki tahapan pembobotan.

Tabel 3.6 Hasil *Text Preprocessing Normalization*

Data Sebelum Normalization	Data Sesudah Normalization
['probolinggo']	['probolinggo']
['wisata', 'alam', 'pegunungan', 'dan', 'air', 'terjun']	['wisata', 'alam', 'pegunungan', 'dan', 'air', 'terjun']
['gunung', 'bromo', 'merupakan', 'gunung', 'berapi', 'aktif', 'di', 'jawa', 'timur', 'yang', 'berada', 'dalam', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'ini', 'dikenal', 'dengan', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'serta', 'kawah', 'aktif', 'yang', 'menjadi', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'beberapa', 'lokasi', 'populer', 'di', 'kawasan', 'ini', 'antara', 'lain', 'kawah', 'bromo', 'pasir', 'berbisik', 'dan', 'bukit', 'teletubbies']	['gunung', 'bromo', 'merupakan', 'gunung', 'berapi', 'aktif', 'di', 'jawa', 'timur', 'yang', 'berada', 'dalam', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'ini', 'dikenal', 'dengan', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'serta', 'kawah', 'aktif', 'yang', 'menjadi', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'beberapa', 'lokasi', 'populer', 'di', 'kawasan', 'ini', 'antara', 'lain', 'kawah', 'bromo', 'pasir', 'berbisik', 'dan', 'bukit', 'teletubbies']

5. Stopword removal

Tahapan ini dilakukan untuk mengeliminasi kata-kata umum (*stopwords*) yang tidak memiliki makna fungsional atau bobot informasi yang signifikan dalam proses analisis teks, seperti kata depan, kata hubung, dan kata ganti.

Tabel 3.7 Hasil *Text Preprocessing Stopword Removal*

Data Sebelum Stopword removal	Data Sesudah Stopword removal
['probolinggo']	['probolinggo']
['wisata', 'alam', 'pegunungan', 'dan', 'air', 'terjun']	['wisata', 'alam', 'pegunungan', 'air', 'terjun']
['gunung', 'bromo', 'merupakan', 'gunung', 'berapi', 'aktif', 'di', 'jawa', 'timur', 'yang', 'berada', 'dalam', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'ini', 'dikenal', 'dengan', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'serta', 'kawah', 'aktif', 'yang', 'menjadi', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'beberapa', 'lokasi', 'populer', 'di', 'kawasan', 'ini', 'antara', 'lain', 'kawah', 'bromo', 'pasir', 'berbisik', 'dan', 'bukit', 'teletubbies']	['gunung', 'bromo', 'gunung', 'berapi', 'aktif', 'jawa', 'timur', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'dikenal', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'kawah', 'aktif', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'lokasi', 'populer', 'kawasan', 'kawah', 'bromo', 'pasir', 'berbisik', 'bukit', 'teletubbies']

6. Stemming

Stemming merupakan tahapan untuk mengubah kata berimbuhan menjadi bentuk dasarnya.

Tabel 3.8 Hasil *Text Preprocessing Stemming*

Data Sebelum <i>Stemming</i>	Data Sesudah <i>Stemming</i>
['probolinggo']	['probolinggo']
['wisata', 'alam', 'pegunungan', 'air', 'terjun']	['wisata', 'alam', 'gunung', 'air', 'terjun']
['gunung', 'bromo', 'gunung', 'berapi', 'aktif', 'jawa', 'timur', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'dikenal', 'lanskap', 'pegunungan', 'vulkanik', 'lautan', 'pasir', 'luas', 'kawah', 'aktif', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'lokasi', 'populer', 'kawasan', 'kawah', 'bromo', 'pasir', 'berbisik', 'bukit', 'teletubbies']	['gunung', 'bromo', 'gunung', 'berapi', 'aktif', 'jawa', 'timur', 'kawasan', 'taman', 'nasional', 'bromo', 'tengger', 'semeru', 'destinasi', 'kenal', 'lanskap', 'gunung', 'vulkanik', 'laut', 'pasir', 'luas', 'kawah', 'aktif', 'daya', 'tarik', 'utama', 'wisata', 'alam', 'lokasi', 'populer', 'kawasan', 'kawah', 'bromo', 'pasir', 'bisik', 'bukit', 'teletubbies']

3.5 Implementasi Metode

Penelitian ini memanfaatkan metode *Content-Based Filtering* yang bertujuan menghasilkan rekomendasi destinasi wisata berdasarkan tingkat kemiripan karakteristik deskripsinya. Representasi dokumen dilakukan melalui pembobotan TF-IDF, yang dilanjutkan dengan kalkulasi derajat kemiripan menggunakan algoritma *Cosine Similarity*. Pendekatan yang sama juga diintegrasikan pada fitur pencarian sistem. Dalam fitur tersebut, TF-IDF digunakan untuk merepresentasikan kata kunci (*query*) pengguna serta konten destinasi yang telah diperkaya dengan ulasan (*review*) wisatawan, sedangkan metode *Cosine Similarity* diterapkan untuk mengukur koefisien kemiripan di antara keduanya.

3.5.1 *Content-Based Filtering*

Metode *Content-Based Filtering* diterapkan supaya menyajikan rekomendasi destinasi wisata berdasarkan tingkat *similarity* karakteristik deskripsi objek wisata. Data yang digunakan merupakan narasi deskripsi destinasi yang telah

melalui tahapan prapemrosesan teks (*text preprocessing*) sehingga hanya menyisakan *term* yang relevan. Setiap dokumen kemudian direpresentasikan ke dalam bentuk matriks vektor memanfaatkan metode TF-IDF. Selanjutnya, derajat kemiripan antar-destinasi dikalkulasi menggunakan algoritma *Cosine Similarity* guna menghasilkan luaran rekomendasi yang sesuai dengan preferensi pengguna.

3.5.1.1 Representasi Dokumen Destinasi Wisata

Pada tahap ini ditampilkan contoh data deskripsi destinasi wisata yang digunakan sebagai dokumen dalam metode *Content-Based Filtering*. Setiap dokumen berisi informasi mengenai karakteristik destinasi wisata yang diperoleh dari berbagai sumber daring. Data tersebut digunakan sebagai dasar dalam proses pengolahan teks dan pembentukan representasi dokumen destinasi wisata.

Tabel 3.9 Contoh Sampel Data Destinasi Wisata

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
1	Gunung Bromo	Probolinggo	Wisata Alam Pegunungan dan Air Terjun	Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Destinasi ini dikenal dengan lanskap pegunungan vulkanik, lautan pasir luas, serta kawah aktif yang menjadi daya tarik utama wisata alam. Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo, Pasir Berbisik, dan Bukit Teletubbies.
2	Kawah Ijen	Banyuwangi	Wisata Alam Pegunungan dan Air Terjun	Kawah Ijen merupakan destinasi wisata alam di Jawa Timur yang terkenal dengan fenomena api biru atau blue fire. Kawasan ini memiliki danau kawah dengan kandungan belerang tinggi serta dikelilingi tebing kawah yang membentuk lanskap vulkanik. Aktivitas penambangan belerang dan jalur pendakian menjadi bagian dari karakter kawasan wisata alam ini.
3	Pantai Papuma	Jember	Wisata Bahari	Pantai Papuma atau Pantai Tanjung Papuma merupakan destinasi wisata bahari di Kabupaten Jember yang memiliki pasir putih, batu karang besar, dan kawasan hutan lindung di sekitarnya. Pantai ini dikenal sebagai lokasi wisata alam dengan panorama laut selatan serta pemandangan matahari terbit dan

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				terbenam. Berbagai aktivitas wisata alam dapat dilakukan di kawasan pantai ini dengan dukungan fasilitas penunjang wisata.

Berdasarkan tabel tersebut, setiap destinasi wisata memiliki informasi berupa nama destinasi, lokasi, kategori wisata, serta deskripsi destinasi. Deskripsi ini memuat karakteristik utama dari setiap destinasi wisata sehingga dapat digunakan sebagai sumber informasi dalam proses analisis konten.

3.5.1.2 Text Preprocessing

Setelah data deskripsi destinasi wisata dikumpulkan, data tersebut melalui tahap *Text Preprocessing* sebagaimana dijelaskan pada Subbab 3.4. Tahapan ini bertujuan untuk membersihkan dan menstandarkan teks sehingga hanya menyisakan kata-kata yang relevan untuk merepresentasikan karakteristik destinasi wisata. Hasil dari proses ini berupa kumpulan kata yang lebih ringkas dan terstruktur, sehingga dapat digunakan sebagai fitur dalam proses pembobotan teks. Contoh hasil data deskripsi destinasi wisata setelah melalui proses *Text Preprocessing* disajikan oleh Tabel 3.10.

Tabel 3.10 Hasil *Text Preprocessing* Data Destinasi Wisata

No	Nama Destinasi	Metadata
1	Gunung Bromo	probolinggo wisata alam gunung air terjun gunung bromo gunung berapi aktif jawa timur kawasan taman nasional bromo tengger semeru destinasi kenal lanskap gunung vulkanik laut pasir luas kawah aktif daya tarik utama wisata alam lokasi populer kawasan kawah bromo pasir bisik bukit teletubbies
2	Kawah Ijen	banyuwangi wisata alam gunung air terjun kawah ijen destinasi wisata alam jawa timur kenal fenomena api biru blue fire kawasan milik danau kawah kandung belerang keliling tebing kawah bentuk lanskap vulkanik aktivitas tambang belerang jalur daki karakter kawasan wisata alam
3	Pantai Papuma	jember wisata bahari pantai papuma pantai tanjung papuma destinasi wisata bahari kabupaten jember milik pasir putih batu karang kawasan hutan lindung pantai kenal lokasi wisata alam panorama laut selatan

No	Nama Destinasi	Metadata
		pandang matahari terbit benam aktivitas wisata alam kawasan pantai dukung fasilitas tunjang wisata

Berdasarkan tabel tersebut, dapat dilihat bahwa teks deskripsi yang semula berupa kalimat panjang telah ditransformasikan menjadi kumpulan *term* yang lebih sederhana, bersih, dan terstruktur. *Term* hasil prapemrosesan tersebut kemudian akan digunakan sebagai input pada tahapan pembobotan dengan menggunakan metode TF-IDF.

3.5.1.3 TF-IDF

Pada tahapan ini, dilakukan proses vektorisasi menggunakan metode TF-IDF untuk mentransformasikan teks deskripsi destinasi wisata ke dalam bentuk numerik agar dapat direpresentasikan sebagai representasi data yang siap diproses lebih lanjut. Penerapan pembobotan ini bertujuan untuk mengukur tingkat signifikansi atau kepentingan suatu *term* di dalam sebuah dokumen spesifik terhadap keseluruhan populasi dokumen dalam korpus.

Dalam penelitian ini, preferensi pengguna direpresentasikan sebagai dokumen preferensi pengguna (D_w), sedangkan dokumen destinasi wisata direpresentasikan sebagai D_1 , D_2 , dan D_3 sebagaimana disajikan oleh Tabel 3.11.

Tabel 3.11 Representasi Dokumen yang Telah Melalui *Text Preprocessing*

No	Dokumen	Metadata
1	Wisatawan (D_w)	wisata alam pandang gunung matahari terbit panorama alam indah
2	Gunung Bromo (D_1)	probolinggo wisata alam gunung air terjun gunung bromo gunung berapi aktif jawa timur kawasan taman nasional bromo tengger semeru destinasi kenal lanskap gunung vulkanik laut pasir luas kawah aktif daya tarik utama wisata alam lokasi populer kawasan kawah bromo pasir bisik bukit teletubbies
3	Kawah Ijen (D_2)	banyuwangi wisata alam gunung air terjun kawah ijen destinasi wisata alam jawa timur kenal fenomena api biru blue fire kawasan milik danau

No	Dokumen	Metadata
		kawah kandung belerang keliling tebing kawah bentuk lanskap vulkanik aktivitas tambang belerang jalur daki karakter kawasan wisata alam
4	Pantai Papuma (D3)	jember wisata bahari pantai papuma pantai tanjung papuma destinasi wisata bahari kabupaten jember milik pasir putih batu karang kawasan hutan lindung pantai kenal lokasi wisata alam panorama laut selatan pandang matahari terbit benam aktivitas wisata alam kawasan pantai dukung fasilitas tunjang wisata

Setelah representasi dokumen terbentuk, selanjutnya dilakukan perhitungan *Term Frequency* (TF) untuk mengetahui frekuensi kemunculan setiap kata pada masing-masing dokumen.

1. *Term Frequency* (TF)

Term Frequency (TF) mengindikasikan frekuensi kemunculan suatu term di dalam sebuah dokumen. Semakin tinggi intensitas kemunculan term tersebut, maka nilai TF yang dihasilkan akan semakin besar. Kalkulasi nilai TF ini dilakukan dengan mengacu pada Persamaan (2.2). Sebagai contoh, perhitungan nilai TF untuk term 'air' pada dokumen D1 dijabarkan sebagai berikut

$$TF(air, D1) = \frac{1}{42} = 0.02381$$

Tabel 3.12 Perhitungan TF

No	Term	Dw	D1	D2	D3	TF Dw	TF D1	TF D2	TF D3
1	air	0	1	1	0	0	0.02381	0.02564	0
2	aktif	0	2	0	0	0	0.04762	0	0
3	aktivitas	0	0	1	1	0	0	0.02564	0.02439
4	alam	2	2	3	2	0.22222	0.04762	0.07692	0.04878
5	api	0	0	1	0	0	0	0.02564	0
...	...	...	...	...	...	...	...	...	...
72	vulkanik	0	1	1	0	0	0.02381	0.02564	0
73	wisata	1	2	3	5	0.11111	0.04762	0.07692	0.12195
TOTAL		9	42	39	41				

Semakin tinggi frekuensi kemunculan suatu kata dalam sebuah dokumen, semakin besar pula nilai TF yang dihasilkan.

2. Inverse Document Frequency (IDF)

IDF diterapkan untuk mengukur tingkat kepentingan suatu *term* dalam seluruh dokumen. Semakin tinggi frekuensi kemunculan kata pada banyak dokumen maka akan memiliki nilai IDF yang lebih rendah. IDF dihitung dengan menggunakan persamaan (2.3). Selanjutnya, nilai IDF dihitung untuk menentukan tingkat kepentingan setiap kata dalam keseluruhan dokumen. Jika kata “air” muncul pada 2 dokumen dari total 4 dokumen, maka:

$$IDF(air) = \log\left(\frac{4}{2}\right) = 0.69315$$

Hasil perhitungan nilai IDF untuk setiap kata pada seluruh dokumen disajikan oleh Tabel 3.13.

Tabel 3.13 Perhitungan IDF

No	Term	IDF
1	Air	0.69315
2	Aktif	1.38629
3	Aktivitas	0.69315
4	Alam	0
5	Api	1.38629
...	...	...
72	Vulkanik	0.69315
73	Wisata	0

Rendahnya nilai IDF dari suatu *term* dipengaruhi oleh tingginya frekuensi kemunculan *term* tersebut di berbagai dokumen. Hal ini terjadi karena *term* yang bersifat umum dianggap tidak memiliki nilai spesifikasi yang kuat dalam mempresentasikan dokumen tertentu.

3. TF-IDF

Bobot TF-IDF diperoleh melalui hasil perkalian antara nilai TF dan IDF. Skor ini merepresentasikan tingkat signifikansi atau kepentingan suatu *term* dalam suatu

dokumen tertentu terhadap seluruh koleksi dokumen lainnya. Kalkulasi bobot TF-IDF ini mengacu pada Persamaan (2.1), di mana contoh perhitungan untuk term 'air' pada dokumen D1 dijabarkan sebagai berikut.

$$TF - IDF(air, D1) = 0.02381 \times 0.69315 = 0.01650$$

Hasil perhitungan TF-IDF untuk setiap kata pada masing-masing dokumen disajikan oleh Tabel 3.14.

Tabel 3.14 Contoh Sampel Data Deskripsi Setelah Perhitungan TF-IDF

No	Term	TF-IDF (Dw)	TF-IDF (D1)	TF-IDF (D2)	TF-IDF (D3)
1	Air	0	0.01650	0.01777	0
2	Aktif	0	0.06627	0	0
3	Aktivitas	0	0	0.01777	0.01691
4	Alam	0	0	0	0
5	Api	0	0	0.03555	0
...	...	...	...	...	...
72	Vulkanik	0	0.01650	0.01777	0
73	Wisata	0	0	0	0

Nilai TF-IDF yang tinggi menunjukkan bahwa suatu *term* memiliki tingkat representasi yang kuat dalam mendeskripsikan karakteristik suatu dokumen, karena *term* tersebut bersifat unik dan jarang muncul pada dokumen lain dalam korpus.

3.5.1.4 Cosine Similarity

Setelah bobot TF-IDF dari setiap dokumen diperoleh, tahapan selanjutnya adalah mengukur nilai kemiripan antara vektor preferensi wisatawan dan dokumen destinasi wisata menggunakan metode *Cosine Similarity*. Perhitungan ini dilakukan guna mengukur tingkat kesesuaian objek wisata tersebut terhadap karakteristik kebutuhan pengguna secara spesifik. Nilai kemiripan dihitung menggunakan *Cosine Similarity* berdasarkan Persamaan (2.4). Hasil perhitungan ini kemudian

digunakan untuk menentukan urutan rekomendasi destinasi wisata dengan tingkat kemiripan tertinggi.

1. Perhitungan *Dot Product*

Langkah awal dalam perhitungan *Cosine Similarity* adalah menghitung nilai *dot product* antara dua vektor dokumen menggunakan persamaan (2.5). Contoh hasil perhitungan *dot product* antara dokumen preferensi pengguna dan masing-masing destinasi wisata dipaparkan pada Tabel 3.15.

Tabel 3.15 Perhitungan *Dot Product*

No	Term	A · B			
		Dw . Dw	Dw . D1	Dw . D2	Dw . D3
1	air	0	0	0	0
2	aktif	0	0	0	0
3	aktivitas	0	0	0	0
4	alam	0	0	0	0
5	api	0	0	0	0
...	...	...	...	...	...
72	vulkanik	0	0	0	0
73	wisata	0	0	0	0
Hasil Dot Product		0.04847	0.00087	0.00023	0.00521

Pada tabel di atas dilakukan perhitungan dot product antara dokumen preferensi pengguna (Dw) dengan masing-masing dokumen destinasi wisata, yaitu Dw dengan Dw, Dw dengan D1, Dw dengan D2, dan Dw dengan D3. Berikut merupakan contoh perhitungan *dot product* antara dokumen Dw dan Dw.

$$\begin{aligned}
 Dw . Dw &= (0 \times 0) + (0 \times 0) + (0 \times 0) + (0 \times 0) + (0.03196 \times 0.03196) \\
 &\quad + (0.154033 \times 0.154033) + \dots + (0.07702 \times 0.07702) \\
 &\quad + (0 \times 0) = 0.04847
 \end{aligned}$$

Berdasarkan hasil perhitungan tersebut diperoleh nilai *dot product* antara dokumen Dw dan Dw sebesar 0.04847.

2. Panjang Vektor (*Norm*)

Tahap berikutnya adalah menghitung panjang vektor atau nilai norma (*norm*) dari masing-masing dokumen. Panjang vektor tersebut diperoleh melalui perhitungan akar kuadrat dari total penjumlahan kuadrat seluruh bobot TF-IDF di dalam dokumen dengan mengacu pada Persamaan (2.6). Adapun contoh kalkulasi panjang vektor untuk dokumen preferensi pengguna (D_w) dijabarkan sebagai berikut.

$$\|D_w\| = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 0.03197^2 + 0.15403^2 + \dots + 0.07702^2 + 0^2}$$

$$\|D_w\| = \sqrt{0 + 0 + 0 + 0 + 0.00102 + 0.02372 + \dots + 0.00593 + 0}$$

$$\|D_w\| = 0.22017$$

Perhitungan yang sama dilakukan pada dokumen destinasi wisata lainnya. Sebagai contoh, perhitungan panjang vektor pada dokumen D1 adalah sebagai berikut.

$$\|D1\| = \sqrt{0.0165^2 + 0.06601^2 + 0^2 + 0^2 + 0.033^2 + 0.033^2 + \dots + 0^2 + 0^2}$$

$$\|D1\| = \sqrt{0.00027 + 0.00436 + 0 + 0 + 0.00108 + 0.00108 + \dots + 0 + 0}$$

$$\|D1\| = 0.18644$$

Hasil perhitungan panjang vektor untuk masing-masing dokumen disajikan pada Tabel 3.16.

Tabel 3.16 Perhitungan Panjang Vektor

No	Term	$\ A\ , \ B\ $			
		$\ D_w\ $	$\ D1\ $	$\ D2\ $	$\ D3\ $
1	air	0	0.00027	0.00031	0
2	aktif	0	0.00436	0	0
3	aktivitas	0	0	0.00031	0.00028
4	alam	0	0	0	0
5	api	0	0	0.00126	0
...	...	...	...	...	...
72	vulkanik	0	0.00027	0.00031	0

No	Term	A . B			
		Dw	D1	D2	D3
73	wisata	0	0	0	0
Hasil Panjang Vektor		0.22017	0.18644	0.17597	0.22034

Nilai panjang vektor ini digunakan sebagai faktor normalisasi dalam perhitungan *similarity*.

3. Hasil Perhitungan *Cosine Similarity*

Setelah nilai dot product dan panjang vektor diperoleh, tahapan terakhir adalah mengalkulasi nilai *Cosine Similarity* dengan memasukkan komponen-komponen tersebut ke dalam persamaan *Cosine Similarity*. Sebagai contoh, perhitungan *Cosine Similarity* antara dokumen preferensi pengguna (Dw) dan dokumen destinasi wisata pertama (D1) dijabarkan sebagai berikut.

$$\text{Cosine Similarity}(Dw, D1) = \frac{0.04847}{0.22017 \times 0.18644} = 0.02134$$

Hasil perhitungan *Cosine Similarity* antara dokumen preferensi pengguna (Dw) dan masing-masing dokumen destinasi wisata disajikan pada Tabel 3.17.

Tabel 3.17 Hasil *Cosine Similarity*

Dokumen	<i>Cosine Similarity</i>
Dw – Dw	1
Dw – D1	0.02134
Dw – D2	0.00608
Dw – D3	0.10736

Hasil tersebut menunjukkan bahwa destinasi Pantai Papuma memiliki tingkat kemiripan paling optimal terhadap preferensi pengguna dibandingkan dengan objek wisata lainnya. Dengan demikian, destinasi ini diprioritaskan oleh sistem sebagai rekomendasi utama.

3.6 Evaluasi Sistem

Evaluasi dilakukan bertujuan untuk mengukur kinerja dari sistem rekomendasi yang telah dikembangkan. Pada penelitian ini, proses pengujian ditinjau dari dua aspek utama, yaitu evaluasi fungsionalitas akurasi rekomendasi dan evaluasi usabilitas (*usability*) dari sisi pengguna. Melalui pendekatan ini, dapat diketahui tingkat relevansi hasil rekomendasi serta parameter kepuasan pengguna terhadap sistem.

3.6.1 Evaluasi Sistem Rekomendasi

Evaluasi sistem rekomendasi dilakukan menggunakan metrik *Precision@K* untuk mengukur tingkat relevansi hasil rekomendasi terhadap preferensi pengguna. Perhitungan dilakukan dengan membandingkan jumlah destinasi wisata yang relevan pada hasil rekomendasi teratas sebanyak K dengan total rekomendasi yang diberikan. Nilai *Precision@K* dihitung menggunakan Persamaan (2.7). Nilai K yang digunakan dalam penelitian ini adalah $K = 5$. Relevansi *item* ditentukan oleh *anotator* berdasarkan kesesuaian destinasi yang direkomendasikan dengan destinasi *input*.

3.6.2 Pengujian Kepuasan Pengguna

Pengukuran tingkat kepuasan pengguna terhadap sistem yang dikembangkan dilakukan melalui evaluasi usabilitas (*usability*). Pengujian ini memanfaatkan metode *Post-Study System Usability Questionnaire* (PSSUQ) sebagai instrumen evaluasi. Instrumen yang digunakan mencakup 16 butir

pertanyaan yang mengadopsi PSSUQ versi 3 (Suwandy et al., 2022). Adapun daftar pertanyaan yang digunakan adalah sebagai berikut:

Tabel 3.18 Daftar Pertanyaan PSSUQ

Aspek	No	Pertanyaan
<i>System Usefulness /</i> Kualitas Sistem (SysUse)	1	Secara keseluruhan, saya puas dengan kemudahan penggunaan sistem ini.
	2	Sistem ini mudah digunakan.
	3	Saya dapat menemukan rekomendasi destinasi wisata dengan cepat menggunakan sistem ini.
	4	Saya merasa nyaman menggunakan sistem ini.
	5	Sistem ini mudah dipelajari.
	6	Sistem ini membantu saya lebih efisien dalam menentukan pilihan destinasi wisata.
<i>Information Quality /</i> Kualitas Informasi (InfoQual)	7	Sistem memberikan pesan atau petunjuk yang jelas ketika terjadi kesalahan.
	8	Jika saya melakukan kesalahan saat menggunakan sistem, saya dapat memperbaikinya dengan mudah.
	9	Informasi yang diberikan sistem jelas dan mudah dipahami.
	10	Saya dapat dengan mudah menemukan informasi yang saya butuhkan.
	11	Informasi yang diberikan sistem membantu saya menentukan pilihan destinasi wisata.
	12	Informasi pada sistem disusun dengan baik dan jelas.
<i>Interface Quality /</i> Kualitas Antar Muka (InterQual)	13	Tampilan antarmuka sistem nyaman dilihat.
	14	Saya menyukai tampilan antarmuka sistem.
	15	Sistem memiliki fitur dan fungsi yang sesuai dengan harapan saya.
<i>Overall</i>	16	Secara keseluruhan, saya puas dengan sistem ini

Kuesioner diberikan kepada responden setelah mereka menggunakan sistem rekomendasi wisata melalui tautan <https://forms.gle/jdkxNTR2vu1Tdixw5> yang disebarkan pada 21 April sampai 24 April 2026. Penilaian terhadap setiap butir pernyataan dilakukan oleh responden dengan memanfaatkan skala Likert pada interval nilai satu sampai dengan tujuh.

3.7 Skenario Uji Coba

Skenario pengujian sistem dirancang untuk mengevaluasi kapabilitas sistem dalam menghasilkan rekomendasi destinasi wisata yang relevan dengan *input* yang diberikan. Pengujian ini dilakukan melalui empat bentuk skenario, yang meliputi

pencarian rekomendasi berdasarkan kategori wisata, pencarian berdasarkan destinasi acuan, serta pencarian berbasis kueri atau kata kunci melalui formulir pencarian.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Terdapat dua *dataset* utama dalam penelitian ini, yaitu *dataset* deskripsi destinasi wisata dan *dataset* ulasan pengguna. *Dataset* destinasi mencakup 27 objek wisata di Jawa Timur dengan atribut nama, lokasi, kategori, dan deskripsi yang digunakan sebagai representasi konten pada sistem *Content-Based Filtering*. Data ini membentuk profil setiap destinasi yang menjadi dasar perhitungan kemiripan antar objek wisata. Contoh sebagian data destinasi dipaparkan oleh Gambar 4.1.

id_destinasi	nama_destinasi	lokasi	kategori	deskripsi	
1	1	gunung bromo	probolinggo	wisata alam pegunungan dan air terjun	Gunung Bromo merupakan gunung berapi aktif di ...
2	2	tumpak sewu	lumajang	wisata alam pegunungan dan air terjun	Air Terjun Tumpak Sewu merupakan destinasi wis...
3	3	jatim park 1	batu	wisata buatan dan rekreasi	Jawa Timur Park 1 merupakan destinasi wisata b...
4	4	pantai papuma	jember	wisata bahari	Pantai Papuma atau Pantai Tanjung Papuma merup...
5	5	baluran	situbondo	wisata hutan dan konservasi alam	Taman Nasional Baluran merupakan kawasan conse...
...	...	...	...	...	...
23	23	coban rondo	malang	wisata alam pegunungan dan air terjun	Coban Rondo merupakan destinasi wisata air ter...
24	24	djawatan	banyuwangi	wisata hutan dan konservasi alam	Djawatan merupakan kawasan wisata hutan di Kab...
25	25	teluk ijo	banyuwangi	wisata bahari	Teluk Ijo merupakan destinasi wisata bahari di...
26	26	kapas biru	lumajang	wisata alam pegunungan dan air terjun	Air Terjun Kapas Biru merupakan destinasi wisa...
27	27	pantai kasap	pacitan	wisata bahari	Pantai Kasap merupakan destinasi wisata bahari...

27 rows x 5 columns

Gambar 4.1 Data Destinasi Wisata

Sedangkan *dataset* ulasan terdiri dari 675 komentar yang telah dibersihkan dan dilabeli menjadi 424 sentimen positif dan 251 sentimen negatif. Data ulasan ini digunakan untuk memperkaya konten dalam fitur pencarian destinasi wisata. Contoh sebagian data ulasan ditunjukkan pada Tabel 4.2.

	id_ulasan	nama_destinasi	teks_ulasan	label
1	1	alas purwo	dulu pernah kesini waktu diklat saka wana bakt...	positif
2	2	alas purwo	Nggak se seram itu loh 🤩 . aku udh pernah kesini...	positif
3	3	alas purwo	Hutan tuh indah banget, sayang sekali terlalu ...	positif
4	4	alas purwo	Sama, aku udh pernah kesitu malah enak bang se...	positif
5	5	alas purwo	saya sering lewat sana malam².. malahan sengaj...	positif
...	...	...	...	...
671	671	tumpak sewu	Tumpak Sewu keren banget, berasa masuk ke duni...	positif
672	672	tumpak sewu	Sudah cantik2 tpi aksesnya cukup buruk. Pemer...	negatif
673	673	tumpak sewu	Tanggal 21 September kemarin saya kesana.. jal...	positif
674	674	tumpak sewu	Alhamdulillah saya udah pernah kesini. Extrim b...	positif
675	675	tumpak sewu	Baru 2minggu lalu kesana, saran sih jgn pake j...	positif

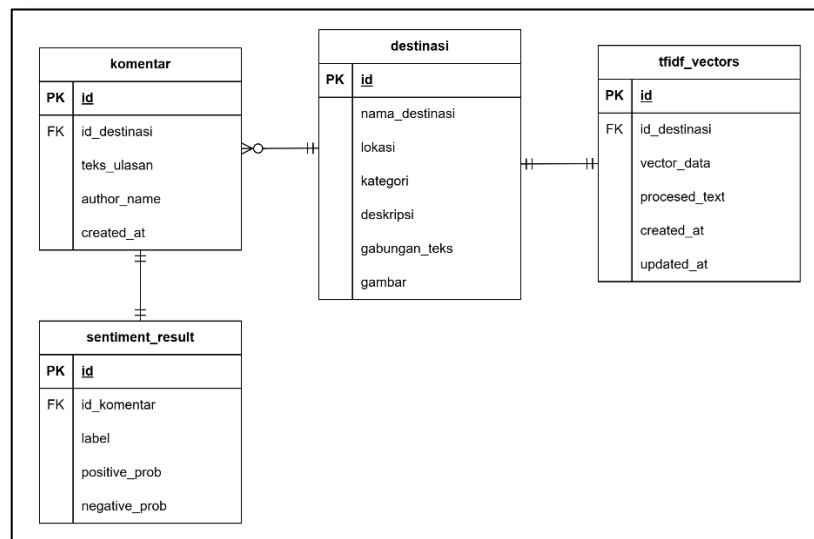
675 rows x 4 columns

Gambar 4.2 Contoh Data Ulasan Pengunjung

Gambar 4.2 memaparkan data ulasan pengunjung yang digunakan. *Dataset* terdiri dari beberapa atribut, yaitu *id_ulasan*, *nama_destinasi*, *teks_ulasan*, dan *label*. Atribut *teks_ulasan* berisi komentar pengguna, sedangkan *label* menunjukkan kategori berdasarkan hasil pelabelan, yaitu positif dan negatif.

4.2 Implementasi Database

Implementasi *database* merupakan tahap penting dalam pembangunan sistem pada penelitian ini. *Database* dirancang menggunakan MySQL sebagai sistem manajemen data yang berperan dalam penyimpanan dan pengelolaan seluruh informasi sistem. *Database* ini terdiri dari beberapa tabel yang saling berelasi untuk mendukung seluruh proses yang berjalan dalam sistem. Representasi struktur basis data dipaparkan pada Gambar 4.3.



Gambar 4.3 Implementasi *Database* Dalam Bentuk ERD

Struktur *database* terdiri dari empat tabel utama, yaitu destinasi, tfidf_vectors, komentar, dan sentiment_results. Tabel destinasi menyimpan informasi lengkap dari 27 destinasi. Tabel *komentar* menyimpan data ulasan pengguna dan berelasi dengan tabel destinasi melalui *foreign key* id_destinasi. Tabel tfidf_vectors menyimpan representasi vektor TF-IDF setiap konten destinasi dan berelasi dengan tabel destinasi melalui *foreign key* id_destinasi. Tabel sentiment_results menyimpan hasil klasifikasi sentimen dari setiap ulasan dan berelasi dengan tabel komentar melalui *foreign key* id_komentar.

4.3 Hasil Sistem Rekomendasi

Bagian ini menyajikan hasil implementasi dan analisis sistem rekomendasi destinasi wisata yang memanfaatkan metode *Content-Based Filtering* dengan pembobotan TF-IDF dan pengukuran kemiripan menggunakan *Cosine Similarity*. Pembahasan dilakukan secara bertahap mulai dari proses *Text Preprocessing*, ekstraksi fitur, perhitungan *similarity* antar destinasi, hingga hasil rekomendasi

yang dihasilkan sistem berdasarkan *input* pengguna. Penyajian hasil pada setiap tahapan bertujuan untuk menunjukkan proses kerja sistem dalam menghasilkan rekomendasi destinasi yang relevan serta menilai kesesuaian hasil rekomendasi berdasarkan karakteristik konten setiap destinasi.

4.3.1 Hasil *Text Preprocessing*

Pada tahap ini, data destinasi wisata digunakan sebagai fitur utama dalam menghitung tingkat kemiripan antar objek wisata. Tahap awal yang dilakukan adalah *Text Preprocessing* untuk membersihkan data dari karakter yang tidak relevan, tanda baca, serta kata-kata yang tidak memiliki makna penting. Tahap ini berfungsi untuk memproses dan kemudian menghasilkan data yang lebih terstruktur untuk kemudian digunakan pada tahap pembobotan fitur.

	nama_destinasi	metadata
0	gunung bromo	probolinggo alam gunung air terjun gunung brom...
1	tumpak sewu	lumajang alam gunung air terjun air terjun tum...
2	jatim park 1	batu buat rekreasi jawa timur park buat kota b...
3	pantai papuma	jember bahari pantai papuma pantai tanjung pap...
4	baluran	situbondo hutan konservasi alam taman nasional...
...	...	...
22	coban rondo	malang alam gunung air terjun coban rondo air ...
23	djawatan	banyuwangi hutan konservasi alam djawatan huta...
24	teluk ijo	banyuwangi bahari teluk ijo bahari kabupaten b...
25	kapas biru	lumajang alam gunung air terjun air terjun kap...
26	pantai kasap	pacitan bahari pantai kasap bahari kabupaten p...

27 rows × 2 columns

Gambar 4.4 Data Destinasi Wisata Setelah Penggabungan Kolom

Setelah proses *Text Preprocessing* selesai, dibuat sebuah kolom baru berupa *metadata* yang merupakan hasil penggabungan dari kolom lokasi, kategori, dan deskripsi. Penggabungan ini bertujuan untuk memperkaya representasi setiap destinasi sehingga sistem dapat menangkap karakteristik konten secara lebih menyeluruh. *Dataset* hasil penggabungan tersebut ditampilkan pada Gambar 4.4.

4.3.2 Perhitungan TF-IDF

Pada sistem rekomendasi yang dikembangkan, data yang digunakan berasal dari *dataset* destinasi wisata yang terdiri dari 27 data destinasi di Jawa Timur. Setiap destinasi memiliki atribut berupa nama destinasi, lokasi, kategori, dan deskripsi. Berdasarkan hasil *Text Preprocessing* sebelumnya, kolom *metadata* digunakan sebagai representasi utama dalam proses rekomendasi. Selanjutnya dilakukan ekstraksi fitur menggunakan metode TF-IDF terhadap kolom *metadata* dengan konfigurasi *ngram_range* = (1,1) dan *max_df*=1.0, sehingga setiap token tunggal (*unigram*) pada seluruh dokumen diikutsertakan dalam perhitungan.

Implementasi perhitungan TF-IDF pada penelitian ini menggunakan *library* *scikit-learn* versi 1.6.1 pada bahasa pemrograman Python. *Library* ini menyediakan dua kelas utama yang digunakan, yaitu *CountVectorizer* untuk perhitungan nilai TF dan *TfidfVectorizer* untuk perhitungan nilai IDF serta TF-IDF dengan normalisasi L2. Nilai TF dihitung menggunakan *CountVectorizer* yang menghasilkan frekuensi kemunculan setiap *term* secara mentah (*raw count*) dalam masing-masing dokumen, tanpa normalisasi. Hasil dari proses perhitungan TF dipaparkan pada Tabel 4.1.

Tabel 4.1 Hasil Perhitungan TF

Id_destinasi	ada	air	ajar	akses	aktif	...	wisata
1	0	1	0	0	2	...	2
2	0	5	0	0	0	...	3
3	0	0	1	0	0	...	3
4	0	0	0	0	0	...	5
5	0	0	0	0	0	...	1
...	...	...	...	...	...	...	...
27	0	1	0	0	0	...	3
df	2	14	1	1	1		27

Nilai IDF dihitung menggunakan formula *smoothed IDF* yang diterapkan oleh scikit-learn sebagai berikut:

$$IDF(air) = \ln\left(\frac{27 + 1}{14 + 1}\right) + 1 = 1.624154$$

Tabel 4.2 Hasil Perhitungan IDF

Term	IDF
ada	3.233592
air	1.624154
ajar	3.639057
akses	3.639057
aktif	3.639057
...	...
wisata	1

Setelah nilai TF dan IDF diperoleh, keduanya dikalikan untuk menghasilkan bobot TF-IDF. Pada tahap awal, nilai TF yang digunakan adalah *raw count* sebagaimana pada Tabel 4.1, sehingga menghasilkan nilai TF-IDF sebelum normalisasi:

$$TF - IDF_{raw}(air, 1) = 1 \times 1.624154 = 1.624154$$

Selanjutnya, hasil perkalian tersebut dinormalisasi menggunakan *L2 norm* (*Euclidean norm*) agar panjang setiap vektor dokumen menjadi 1. Nilai penyebut pada persamaan tersebut ditentukan berdasarkan akar dari jumlah kuadrat

keseluruhan bobot *term* pada dokumen terkait. Contoh perhitungan normalisasi untuk term "air" pada destinasi 1 adalah:

$$TF - IDF_{l2\ norm}(air, 1) = \frac{1.624154}{\sqrt{0^2 + 1.62415^2 + \dots + 2^2}} = \frac{1.624154}{20.935} = 0.077758$$

Tabel 4.3 Hasil Perhitungan TF-IDF

Id_destinasi	ada	air	ajar	akses	aktif	...	wisata
1	0	0.07758	0	0	0.34765	...	0.09553
2	0	0.40940	0	0	0	...	0.15124
3	0	0	0.20603	0	0	...	0.16985
4	0	0	0	0	0	...	0.24555
5	0	0	0	0	0	...	0.04838
...	...	...	...	...	...	...	...
27	0	0.09860	0	0	0	...	0.18212

Nilai TF-IDF yang lebih tinggi menunjukkan tingkat representasi yang lebih kuat dari suatu *term* dalam mendeskripsikan karakteristik objek wisata dibanding *term* lain. Perbedaan bobot yang dihasilkan mengonfirmasi kemampuan metode TF-IDF dalam mengekstrak fitur relevan dan membedakan tiap-tiap destinasi, yang pada gilirannya akan mendukung efektivitas komputasi kemiripan pada langkah berikutnya.

4.3.3 Perhitungan *Cosine Similarity*

Pengukuran tingkat kemiripan antar-destinasi pada tahapan berikutnya dilakukan melalui pendekatan *Cosine Similarity*. Pendekatan ini bekerja dengan menghitung sudut antar vektor fitur sehingga menghasilkan nilai kemiripan antara 0 hingga 1. Nilai yang semakin mendekati angka 1 menunjukkan kedua destinasi memiliki tingkat kemiripan yang tinggi.

Karena implementasi normalisasi L2 telah dilakukan pada tahap sebelumnya, penyebut dalam rumus *Cosine Similarity* memiliki nilai sama dengan

1. Dengan demikian, perhitungan *Cosine Similarity* disederhanakan menjadi perkalian titik (*dot product*) antar vektor fitur. Contoh perhitungan kemiripan antara destinasi 1 dan destinasi 2 adalah sebagai berikut:

$$\begin{aligned} Sim(d1, d2) &= (0 \times 0) + (0.07758 \times 0.40940) + \dots + (0.09553 \times 0.15124) \\ &= 0.2409 \end{aligned}$$

Hasil perhitungan secara menyeluruh untuk seluruh pasangan destinasi wisata ditunjukkan pada matriks *Cosine Similarity* di Tabel 4.4.

Tabel 4.4 Perhitungan *Cosine Similarity*

id	1	2	3	4	...	24	25	26	27
1	1	0.2409	0.0850	0.1160	...	0.0514	0.1343	0.2541	0.1193
2	0.2409	1	0.0521	0.1000	...	0.1186	0.1330	0.6356	0.1377
3	0.0850	0.0521	1	0.0941	...	0.0679	0.0711	0.0467	0.0519
4	0.1160	0.1000	0.0941	1	...	0.1448	0.2207	0.0796	0.3854
...	...	...	...	...	...	...	...	...	...
24	0.0514	0.1186	0.0679	0.1448	...	1	0.1812	0.0824	0.0868
25	0.1343	0.1330	0.0711	0.2207	...	0.1812	1	0.1855	0.3188
26	0.2541	0.6356	0.0467	0.0796	...	0.0824	0.1854	1	0.1685
27	0.1193	0.1378	0.0519	0.3854	...	0.0868	0.3188	0.1685	1

Tabel 4.4 menampilkan matriks kedekatan antar seluruh destinasi wisata. Pada matriks tersebut, nilai 1 pada garis diagonal menunjukkan hasil perhitungan kemiripan suatu destinasi terhadap dirinya sendiri. Secara umum, semakin besar nilai *similarity* antar dua destinasi, maka tingkat kemiripannya semakin tinggi. Sistem selanjutnya menggunakan matriks nilai kemiripan ini sebagai dasar pengurutan destinasi dari skor tertinggi ke terendah. Pada tahap akhir, sistem menarik destinasi dengan skor kemiripan tertinggi untuk menjadi daftar rekomendasi (*Top-N Recommendation*) bagi pengguna.

4.3.4 Hasil Rekomendasi

Setelah seluruh data destinasi wisata melalui proses pembobotan menggunakan TF-IDF dan perhitungan *similarity* menggunakan *Cosine Similarity*, tahap berikutnya adalah menghasilkan rekomendasi berdasarkan *input* pengguna. *Input* yang digunakan berupa pemilihan salah satu destinasi wisata yang tersedia pada sistem, seperti ditunjukkan oleh Tabel 4.5.

Tabel 4.5 *Input* Data

Nama Destinasi	Lokasi	Kategori	Deskripsi
Tumpak Sewu	Lumajang	Wisata alam pegunungan dan air terjun	Air Terjun Tumpak Sewu merupakan destinasi wisata alam yang terletak di wilayah perbatasan Kabupaten Lumajang dan Kabupaten Malang. Air terjun ini dikenal memiliki aliran air yang menyerupai tirai dengan latar tebing tinggi dan lingkungan alam pegunungan. Kawasan ini menjadi daya tarik wisata alam karena keindahan lanskap dan karakter air terjunnya yang unik.

Berdasarkan *input* tersebut, sistem akan mencocokkan deskripsi destinasi yang dipilih dengan deskripsi destinasi lainnya dalam *dataset*. Proses pencocokan dilakukan dengan memanfaatkan matriks *Cosine Similarity* yang telah dihitung sebelumnya untuk menentukan tingkat kemiripan antar destinasi. Pada halaman rincian destinasi yang dipilih, sistem secara otomatis menyajikan daftar rekomendasi objek wisata lain yang memiliki nilai kemiripan tertinggi. Rekomendasi tersebut ditempatkan pada bagian bawah halaman guna memudahkan pengguna dalam menemukan destinasi alternatif yang relevan tanpa harus melakukan proses pencarian ulang.

Proses rekomendasi dilakukan dengan mengidentifikasi destinasi yang dipilih, kemudian mengambil nilai kemiripan antara destinasi tersebut dengan seluruh destinasi lainnya. Tahap berikutnya adalah melakukan pengurutan nilai dari

yang terbesar hingga terkecil, di mana beberapa objek wisata dengan skor kemiripan tertinggi akan dipilih dan ditampilkan sebagai hasil akhir rekomendasi.

Hasil rekomendasi yang dihasilkan berdasarkan *input* yang dipilih ditampilkan pada

Tabel 4.6.

Tabel 4.6 Hasil Rekomendasi

id	Nama Destinasi	Lokasi	Kategori	Deskripsi	Skor Similarity
26	Kapas Biru	Lumajang	Wisata alam pegunungan dan air terjun	Air Terjun Kapas Biru merupakan destinasi wisata alam di Kabupaten Lumajang yang memiliki aliran air bertingkat dengan latar tebing curam. Air terjun ini berada di kawasan lembah sungai yang berhulu di lereng Gunung Semeru dan dikelilingi lanskap alam pegunungan. Karakter air terjun bertingkat dan lingkungan alamnya menjadi ciri utama kawasan wisata ini.	0.636
21	Madakaripura	Probolinggo	wisata alam pegunungan dan air terjun	Air Terjun Madakaripura merupakan destinasi wisata alam di Kabupaten Probolinggo yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Air terjun ini memiliki ketinggian dan aliran air yang mengelilingi dinding tebing lembah sempit. Lanskap air terjun dan lingkungan pegunungan menjadi karakter utama kawasan wisata ini.	0.559
23	Coban Rondo	Malang	wisata alam pegunungan dan air terjun	Coban Rondo merupakan destinasi wisata air terjun di Kabupaten Malang yang berada di kawasan pegunungan dengan lingkungan alam yang sejuk. Air terjun ini dikenal memiliki akses yang relatif mudah serta dikelilingi area hijau yang masih terjaga. Karakter kawasan ini menjadikannya sebagai salah satu destinasi wisata alam pegunungan di Malang.	0.519

id	Nama Destinasi	Lokasi	Kategori	Deskripsi	Skor Similarity
10	Coban Talun	Batu	wisata alam pegunungan dan air terjun	Coban Talun merupakan destinasi wisata air terjun di Kota Batu yang berada di kawasan pegunungan dengan lingkungan alam yang sejuk. Air terjun ini memiliki aliran air jernih dan dikelilingi kawasan hutan pegunungan. Karakter alamnya menjadikan Coban Talun sebagai destinasi wisata alam pegunungan.	0.507
16	Ranu Kumbolo	Lumajang	wisata alam pegunungan dan air terjun	Ranu Kumbolo merupakan danau air tawar yang berada di kawasan Taman Nasional Bromo Tengger Semeru. Danau ini terletak di kaki Gunung Semeru dan dikelilingi lanskap perbukitan serta hutan pegunungan. Kondisi alam yang masih terjaga menjadikan Ranu Kumbolo sebagai bagian dari wisata alam pegunungan.	0.335

Berdasarkan hasil rekomendasi yang diperoleh, dapat dianalisis bahwa destinasi yang dihasilkan sistem memiliki karakteristik yang serupa dengan destinasi acuan, khususnya pada aspek kategori dan kesamaan deskripsi. Nilai *similarity* tertinggi sebesar 0,636 antara Tumpak Sewu dan Kapas Biru menunjukkan kedua destinasi memiliki kedekatan karakteristik yang tinggi. Hal ini dipengaruhi oleh kesamaan kategori, yaitu wisata alam pegunungan dan air terjun, serta kemiripan deskripsi yang sama-sama merepresentasikan unsur air terjun, tebing, dan lanskap pegunungan. Hasil ini menunjukkan bahwa perhitungan *Cosine Similarity* mampu mengidentifikasi hubungan relevan antar destinasi berdasarkan kesamaan konten.

4.3.5 Hasil Rekomendasi Berdasarkan Pencarian Kata Kunci

Selain memberikan rekomendasi berdasarkan kedekatan antar objek wisata (*item-to-item*), sistem yang dikembangkan juga menyediakan fitur pencarian berbasis kata kunci bebas (*query-to-item*). Pada fitur ini, kueri yang dimasukkan oleh pengguna diproses dan diekstraksi fitur teksnya sebagai dokumen baru. Dokumen kueri tersebut kemudian dihitung tingkat kemiripannya terhadap seluruh dokumen destinasi wisata di dalam *database*. Untuk menguji keandalan dan akurasi fitur ini, dilakukan pengujian sistem menggunakan Skenario 3 dengan menginputkan kata kunci pencarian "Pantai bersih".

4.3.5.1 Perhitungan TF-IDF dan *Cosine Similarity*

Untuk membuktikan validitas matematika dari algoritma yang tertanam pada sistem pencarian, dijabarkan proses perhitungan manual mulai dari pembobotan TF-IDF kueri hingga pengukuran *Cosine Similarity* terhadap lima sampel destinasi rekomendasi teratas.

Sebagai catatan pembatasan ruang vektor, mengingat jumlah dimensi kosa kata unik (*vocabulary*) pada korpus data mencapai ratusan kata yang tidak memungkinkan untuk dijabarkan seluruhnya secara tertulis, simulasi perhitungan manual ini menggunakan notasi agregasi *Sigma* (Σ) untuk merepresentasikan nilai dimensi kata lain di luar kata pencarian. Hal ini bertujuan agar alur logika perhitungan dan validasi akhir algoritma tetap dapat dibuktikan secara matematis, tanpa merubah nilai panjang vektor atau *Euclidean Norm* ($\|D\|$) dari keadaan sistem yang sebenarnya.

1. Perhitungan *Inverse Document Frequency* (IDF)

Pada skenario pencarian ini, jumlah keseluruhan dokumen (N) adalah 28 dokumen, yang terdiri dari 27 dokumen destinasi wisata dan 1 dokumen kueri pencarian. Berdasarkan hasil ekstraksi, kata "pantai" muncul pada 14 dokumen, sedangkan kata "bersih" muncul pada 6 dokumen. Perhitungan nilai *Smoothed* IDF untuk kedua kata tersebut adalah sebagai berikut:

$$IDF_{pantai} = \ln\left(\frac{1+28}{1+14}\right) + 1 = \ln\left(\frac{29}{15}\right) + 1 = \ln(1.9333) + 1 = 0.6592 + 1$$

$$IDF_{pantai} = 1.6592$$

$$IDF_{bersih} = \ln\left(\frac{1+28}{1+6}\right) + 1 = \ln\left(\frac{29}{7}\right) + 1 = \ln(4.1428) + 1 = 1.4214 + 1$$

$$IDF_{bersih} = 2.4214$$

2. Pembobotan dan Normalisasi Vektor Kueri (Q)

Pada dokumen kueri ("Pantai bersih"), nilai *Term Frequency* (TF) untuk kata "pantai" dan "bersih" masing-masing adalah 1. Bobot awal dihitung dengan mengalikan nilai TF dan IDF:

$$w_{q,pantai} = TF_{q,pantai} \times IDF_{pantai} = 1 \times 1.6592 = 1.6592$$

$$w_{q,bersih} = TF_{q,bersih} \times IDF_{bersih} = 1 \times 2.4214 = 2.4214$$

Untuk mendapatkan bobot akhir, dilakukan normalisasi *L2 Norm* dengan membagi bobot awal dengan panjang vektor kueri ($\|Q\|$):

$$\|Q\| = \sqrt{w_{q,pantai} + w_{q,bersih}}$$

$$\|Q\| = \sqrt{(1.6592)^2 + (2.4214)^2} = \sqrt{2.7529 + 5.8631} = \sqrt{8.6160}$$

$$\|Q\| = 2.9353$$

Maka, nilai bobot kueri ternormalisasi (Q_{norm}) adalah:

$$TF - IDF_{\text{query}, \text{pantai}} = \frac{w_{q, \text{pantai}}}{\|Q\|} = \frac{1.6592}{2.9353} = 0.5653$$

$$TF - IDF_{\text{query}, \text{bersih}} = \frac{w_{q, \text{bersih}}}{\|Q\|} = \frac{2.4214}{2.9353} = 0.8249$$

3. Pembobotan dan *Cosine Similarity* 5 Destinasi Rekomendasi Teratas

Setelah vektor kueri terbentuk, sistem mengukur kedekatan (*Cosine Similarity*) antara kueri dengan dokumen destinasi menggunakan perkalian titik (*Dot Product*). Berikut adalah perhitungannya terhadap 5 destinasi rekomendasi teratas.

a. Dokumen Pantai 3 Warna

Pada dokumen ini, kata "pantai" muncul sebanyak 25 kali, sedangkan kata "bersih" muncul 1 kali. Bobot awal TF-IDF dihitung dengan mengalikan frekuensi kemunculan dengan nilai IDF kueri.

$$w_{\text{pantai3warna}, \text{pantai}} = TF_{p3\text{warna}, \text{pantai}} \times IDF_{\text{pantai}} = 25 \times 1.6592 = 41.4800$$

$$w_{\text{pantai3warna}, \text{bersih}} = TF_{p3\text{warna}, \text{bersih}} \times IDF_{\text{pantai}} = 1 \times 2.4214 = 2.4214$$

Selanjutnya, dilakukan normalisasi L2 untuk mendapatkan bobot akhir. Diketahui bahwa total kuadrat dari kata lainnya dalam dokumen ini adalah 1387.5101, sehingga perhitungan panjang vektor dan bobot ternormalisasinya adalah sebagai berikut:

$$\|D\| = \sqrt{(w_{d, \text{pantai}})^2 + (w_{d, \text{bersih}})^2 + \sum (w_{d, \text{kata_lain}})^2}$$

$$\|D\| = \sqrt{(41.4800)^2 + (2.4214)^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{1720.5904 + 5.8631 + 1387.5101}$$

$$\|D\| = \sqrt{3113.9636}$$

$$\|D\| = 55.8029$$

$$TF - IDF_{pantai3warna,pantai} = \frac{41.4800}{55.8029} = 0.7433$$

$$TF - IDF_{pantai3warna,bersih} = \frac{2.4214}{55.8029} = 0.0434$$

Setelah bobot ternormalisasi didapatkan, nilai *Cosine Similarity* dihitung dengan mengalikan bobot dokumen tersebut dengan bobot kueri pencarian secara berpasangan:

$$CosineSimilarity(Q, D_{pantai3warna}) = Q \cdot D$$

$$CosineSimilarity(Q, D) = (0.5653 \times 0.7433) + (0.8249 \times 0.0434) + 0$$

$$CosineSimilarity(Q, D_{pantai3warna}) = 0.42018 + 0.03580 + 0$$

$$CosineSimilarity(Q, D_{pantai3warna}) = 0.45598$$

b. Dokumen Balekambang

Berdasarkan data destinasi, kata "pantai" muncul sebanyak 15 kali dan kata "bersih" muncul 1 kali, sehingga perhitungan bobot awalnya adalah:

$$w_{balekambang,pantai} = TF_{balekambang,pantai} \times IDF_{pantai} = 15 \times 1.6592$$

$$w_{balekambang,pantai} = 24.8880$$

$$w_{balekambang,bersih} = TF_{balekambang,bersih} \times IDF_{bersih} = 1 \times 2.4214 = 2.4214$$

Untuk menormalisasi vektor tersebut, diketahui total kuadrat dari kata lainnya sebesar 632.6395. Perhitungan panjang vektor dan pembagian bobotnya dilakukan sebagai berikut:

$$\|D\| = \sqrt{(w_{d,pantai})^2 + (w_{d,bersih})^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{(24.8880)^2 + (2.4214)^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{619.4125 + 5.8631 + 632.6395}$$

$$\|D\| = \sqrt{1257.9152}$$

$$\|D\| = 35.4671$$

$$TF - IDF_{d,t} = \frac{w_{d,t}}{\|D\|}$$

$$TF - IDF_{balekambang,pantai} = \frac{24.8880}{35.4671} = 0.7017$$

$$TF - IDF_{pantai3warna,bersih} = \frac{2.4214}{35.4671} = 0.0683$$

Perhitungan nilai *Cosine Similarity* untuk dokumen Balekambang diselesaikan melalui operasi penjumlahan perkalian titik:

$$CosineSimilarity(Q, D_{balekambang}) = Q \cdot D$$

$$CosineSimilarity(Q, D) = (0.5653 \times 0.7017) + (0.8249 \times 0.0683) + 0$$

$$CosineSimilarity(Q, D_{balekambang}) = 0.3967 + 0.0563 + 0$$

$$CosineSimilarity(Q, D_{balekambang}) = 0.4530$$

c. Dokumen Pantai Klayar

Pada dokumen Pantai Klayar, kata "pantai" memiliki frekuensi sebanyak 18 kali, namun kata "bersih" tidak ditemukan (0 kali). Nilai bobot awalnya adalah:

$$w_{klayar,pantai} = TF_{klayar,pantai} \times IDF_{pantai} = 18 \times 1.6592 = 29.8656$$

$$w_{klayar,bersih} = TF_{klayar,bersih} \times IDF_{bersih} = 0 \times 2.4214 = 0$$

Proses normalisasi menggunakan total kuadrat kata lainnya sebesar 972.4720, yang menghasilkan perhitungan bobot akhir sebagai berikut:

$$\|D\| = \sqrt{(w_{d,pantai})^2 + (w_{d,bersih})^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{(29.8656)^2 + (0)^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{891.9540 + 0 + 972.4720}$$

$$\|D\| = \sqrt{1864.4260}$$

$$\|D\| = 43.1790$$

$$TF - IDF_{d,t} = \frac{w_{d,t}}{\|D\|}$$

$$TF - IDF_{klayar,pantai} = \frac{29.8656}{43.1790} = 0.6917$$

$$TF - IDF_{klayar,bersih} = \frac{0}{43.1790} = 0$$

Karena salah satu kata tidak muncul, perkalian *Cosine Similarity* pada dokumen ini sebagian tereliminasi menjadi 0:

$$CosineSimilarity(Q, D_{klayar}) = Q \cdot D$$

$$\text{CosineSimilarity}(Q, D_{klayar}) = (0.5653 \times 0.6917) + (0.8249 \times 0) + 0$$

$$\text{CosineSimilarity}(Q, D_{klayar}) = 0.3910 + 0 + 0$$

$$\text{CosineSimilarity}(Q, D_{klayar}) = 0.3910$$

d. Dokumen Teluk Asmara

Kata "pantai" pada dokumen ini muncul sebanyak 16 kali dan "bersih" muncul 1 kali. Pembobotan mentah untuk kedua kata tersebut adalah:

$$w_{telukasmara,pantai} = TF_{telukasmara,pantai} \times IDF_{pantai} = 16 \times 1.6592$$

$$w_{telukasmara,pantai} = 26.5472$$

$$w_{telukasmara,bersih} = TF_{telukasmara,bersih} \times IDF_{bersih} = 1 \times 2.4214 = 2.4214$$

Dengan nilai akumulasi kuadrat kata lainnya sebesar 2113.1383, perhitungan panjang vektor dan normalisasi bobot dilakukan sebagai berikut:

$$\|D\| = \sqrt{(w_{d,pantai})^2 + (w_{d,bersih})^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{(26.5472)^2 + (2.4214)^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{704.7538 + 5.8632 + 2113.1383}$$

$$\|D\| = \sqrt{2823.7553}$$

$$\|D\| = 53.1390$$

$$TF - IDF_{d,t} = \frac{w_{d,t}}{\|D\|}$$

$$TF - IDF_{telukasmara,pantai} = \frac{26.5472}{53.1390} = 0.4996$$

$$TF - IDF_{telukasmara,bersih} = \frac{2.4214}{53.1390} = 0.0456$$

Selanjutnya, tingkat kedekatan dokumen ini diukur menggunakan *Cosine Similarity* yang menghasilkan:

$$CosineSimilarity(Q, D_{telukasmara}) = Q \cdot D$$

$$CosineSimilarity(Q, D) = (0.5653 \times 0.4996) + (0.8249 \times 0.0456) + 0$$

$$CosineSimilarity(Q, D_{telukasmara}) = 0.2824 + 0.0376 + 0$$

$$CosineSimilarity(Q, D_{telukasmara}) = 0.3200$$

e. Dokumen Pantai Kasap

Kata "pantai" muncul sebanyak 14 kali dan kata "bersih" muncul 1 kali pada dokumen ini, sehingga menghasilkan bobot awal:

$$w_{pantaikasap,pantai} = TF_{d,pantai} \times IDF_{pantai} = 14 \times 1.6592 = 23.2288$$

$$w_{pantaikasap,bersih} = TF_{d,bersih} \times IDF_{bersih} = 1 \times 2.4214 = 2.4214$$

Diketahui total kuadrat kata lainnya bernilai 1947.4946. Nilai ini digunakan untuk mencari panjang vektor dan bobot ternormalisasi:

$$\|D\| = \sqrt{(w_{d,pantai})^2 + (w_{d,bersih})^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{(23.2288)^2 + (2.4214)^2 + \sum (w_{d,kata_lain})^2}$$

$$\|D\| = \sqrt{539.5771 + 5.8632 + 1947.4946}$$

$$\|D\| = \sqrt{2492.9349}$$

$$\|D\| = 49.9293$$

$$TF - IDF_{d,t} = \frac{w_{d,t}}{\|D\|}$$

$$TF - IDF_{pantaikasap,pantai} = \frac{23.2288}{49.9293} = 0.4652$$

$$TF - IDF_{pantaikasap,bersih} = \frac{2.4214}{49.9293} = 0.0485$$

Tahap terakhir, nilai *Cosine Similarity* dihitung dengan menjumlahkan perkalian pasangan bobot tersebut:

$$CosineSimilarity(Q, D_{pantaikasap}) = Q \cdot D$$

$$CosineSimilarity(Q, D) = (0.5653 \times 0.4652) + (0.8249 \times 0.0485) + 0$$

$$CosineSimilarity(Q, D_{pantaikasap}) = 0.2630 + 0.0400 + 0$$

$$CosineSimilarity(Q, D_{pantaikasap}) = 0.3030$$

4.3.5.2 Hasil Rekomendasi Pencarian

Setelah nilai *Cosine Similarity* antara kueri dan seluruh destinasi diperoleh, nilai tersebut selanjutnya diurutkan dari yang nilai tertinggi ke terendah. Lima destinasi dengan nilai kemiripan tertinggi kemudian ditampilkan kepada pengguna sebagai hasil rekomendasi (*Top-5 Recommendation*). Hasil rekomendasi yang dihasilkan berdasarkan pencarian kata kunci "Pantai bersih" ditampilkan pada Tabel 4.7.

Tabel 4.7 Hasil Rekomendasi Skenario 3 (Kueri “Pantai bersih”)

No	Destinasi	Skor Similarity
1	Pantai Tiga Warna	0.456
2	Balekambang	0.435
3	Pantai Klayar	0.391
4	Teluk Asmara	0.320
5	Pantai Kasap	0.303

Berdasarkan Tabel 4.7, hasil rekomendasi sistem sudah sesuai dengan perhitungan manual yang dijabarkan pada sub-bab sebelumnya. Destinasi Pantai Tiga Warna dan Balekambang menempati peringkat teratas karena kata "pantai" dan "bersih" sering muncul pada data destinasinya, sehingga menghasilkan nilai kemiripan tertinggi yang mendekati angka 1. Hasil ini menunjukkan bahwa fungsionalitas fitur pencarian pada sistem telah berjalan secara optimal dan presisi, selaras dengan mekanisme metode yang dirancang.

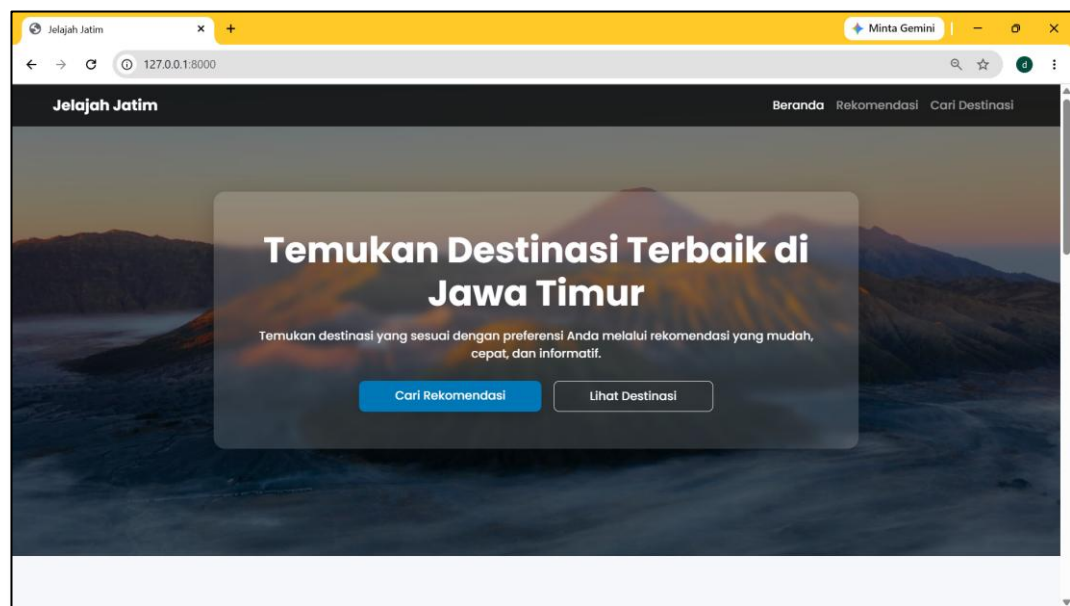
4.4 Implementasi Sistem

Sistem rekomendasi diimplementasikan dalam bentuk web berbasis *framework* Laravel dengan bahasa pemrograman PHP. Pendekatan *Content-Based Filtering* diterapkan supaya menyajikan rekomendasi destinasi wisata berdasarkan nilai *similarity* deskripsi melalui perhitungan TF-IDF dan *Cosine Similarity*. Disisi lain, sistem ini menyediakan fitur pencarian yang memanfaatkan data ulasan wisatawan, di mana proses komputasinya diintegrasikan dengan layanan FastAPI menggunakan metode TF-IDF dan *Cosine Similarity*.

4.4.1 Halaman Beranda

Halaman beranda adalah antarmuka utama yang pertama kali disajikan kepada pengguna ketika pertamakali mengakses sistem. Pada halaman ini terdapat

bilah navigasi (*navbar*) yang berfungsi untuk mengarahkan pengguna ke berbagai halaman lainnya. Selain itu, halaman beranda juga menyajikan informasi singkat terkait sistem, serta menyediakan tombol ‘Cari Rekomendasi’ untuk mengakses halaman pencarian dan tombol ‘Lihat Destinasi’ untuk menuju ke halaman seluruh destinasi. Halaman beranda ditampilkan pada Gambar 4.5.

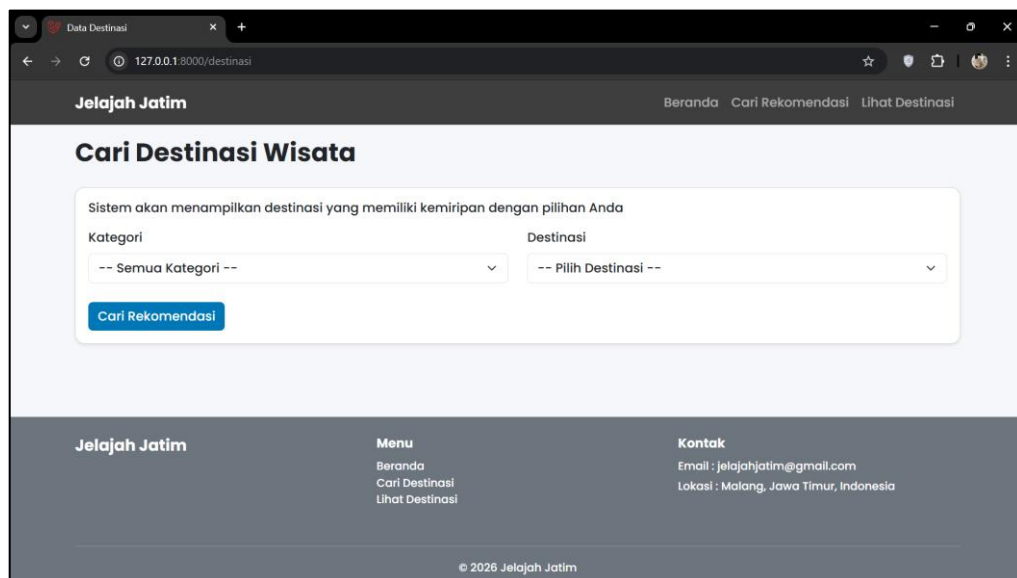


Gambar 4.5 Tampilan Halaman Beranda

Gambar 4.5 menampilkan halaman *home* dari sistem Jelajah Jatim dengan latar belakang foto pemandangan Gunung Bromo, Jawa Timur, pada bagian depan dan dilengkapi dengan judul serta deskripsi singkat sistem sebagai pengantar bagi pengguna yang baru mengakses sistem. Pada *navbar* terdapat tombol ‘Cari Destinasi’ yang kemudian akan diarahkan ke halaman *searching* dimana pengguna menginputkan kata kunci sesuai dengan preferensi.

4.4.2 Halaman Cari Rekomendasi

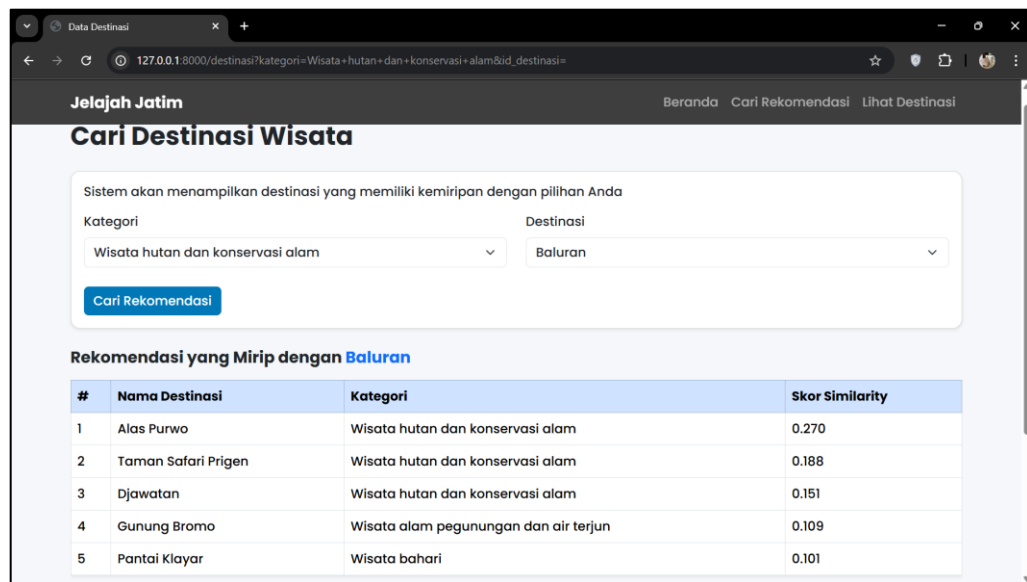
Halaman cari rekomendasi memungkinkan pengguna untuk mencari destinasi wisata yang memiliki kemiripan dengan destinasi yang dipilih. Halaman ini terdiri dari dua *dropdown*, yaitu *dropdown* kategori dan *dropdown* destinasi wisata. *Dropdown* kategori berfungsi untuk memfilter daftar destinasi berdasarkan jenis wisatanya, sehingga pengguna dapat lebih mudah menemukan destinasi yang diinginkan. Setelah pengguna memilih kategori dan destinasi yang diinginkan, pengguna dapat menekan tombol ‘Cari Rekomendasi’ untuk memperoleh daftar destinasi wisata yang paling relevan. Tampilan halaman cari rekomendasi ditampilkan pada Gambar 4.6.



Gambar 4.6 Tampilan Halaman Cari Rekomendasi

Gambar 4.6 menampilkan halaman cari rekomendasi sebelum pengguna memasukkan *input* pencarian. Dalam kondisi ini, *dropdown* kategori menampilkan pilihan ‘Semua Kategori’ dan *dropdown* destinasi menampilkan pilihan ‘Pilih

Destinasi’, serta terdapat tombol ‘Cari Rekomendasi’ yang siap digunakan oleh pengguna.



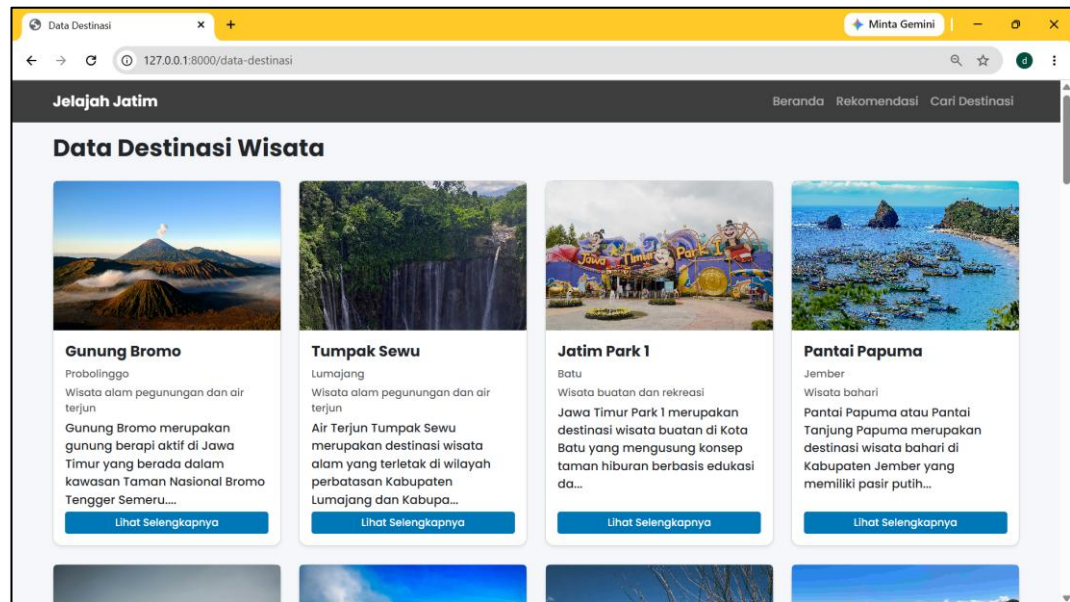
Gambar 4.7 Tampilan Halaman Cari Rekomendasi dengan Hasil Rekomendasi

Gambar 4.7 memperlihatkan halaman hasil pencarian rekomendasi. Setelah tombol ‘Cari Rekomendasi’ ditekan, sistem memproses data untuk menampilkan lima objek wisata yang memiliki karakteristik paling serupa berdasarkan perolehan skor *similarity* tertinggi. Hasil rekomendasi ditampilkan dalam bentuk tabel di bawah *form* pencarian yang memuat informasi nomor urut, nama destinasi, kategori dan skor *similarity*.

4.4.3 Halaman Seluruh Data Destinasi

Halaman seluruh data destinasi digunakan dalam menyajikan daftar lengkap objek wisata yang tersedia di dalam sistem. Setiap destinasi ditampilkan memuat foto, nama destinasi, lokasi, kategori, serta deskripsi singkat. Pengguna dapat menekan tombol ‘Lihat Selengkapnya’ pada masing-masing destinasi untuk

mengakses detail informasi lebih lengkap mengenai destinasi tersebut. Halaman ini ditampilkan pada Gambar 4.8.



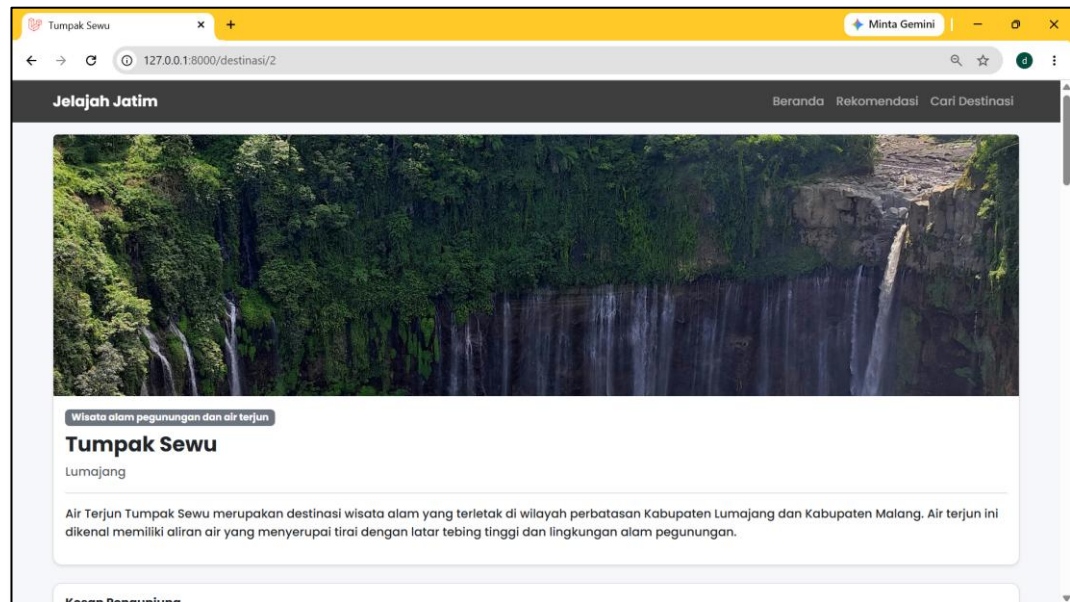
Gambar 4.8 Tampilan Halaman Seluruh Data Destinasi

Gambar 4.8 menunjukkan tampilan halaman seluruh data destinasi yang menyajikan daftar destinasi wisata yang memuat foto, nama destinasi, lokasi, kategori, serta deskripsi singkat yang dapat membantu pengguna dalam mengenal destinasi sebelum melihat informasi lebih lengkap.

4.4.4 Halaman Rincian Destinasi

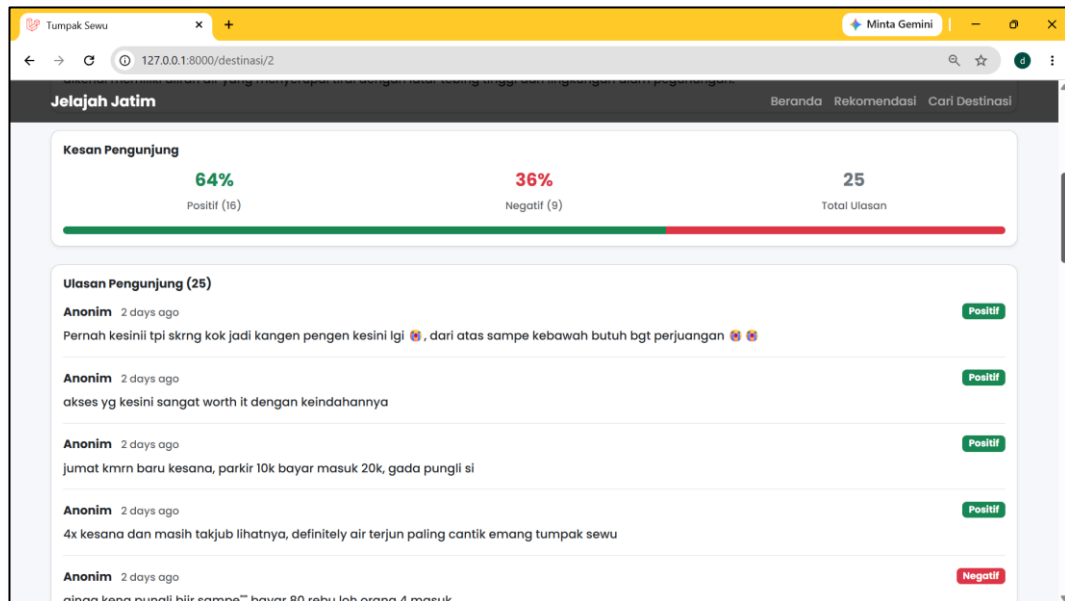
Halaman rincian destinasi berfungsi menyajikan data komprehensif mengenai objek wisata yang telah dipilih oleh pengguna. Informasi yang termuat di halaman rincian destinasi ini meliputi nama destinasi, lokasi, kategori, deskripsi,

kesan pengunjung, ulasan wisatawan, serta rekomendasi objek wisata lain yang relevan. Halaman rincian destinasi ditampilkan pada Gambar 4.9.



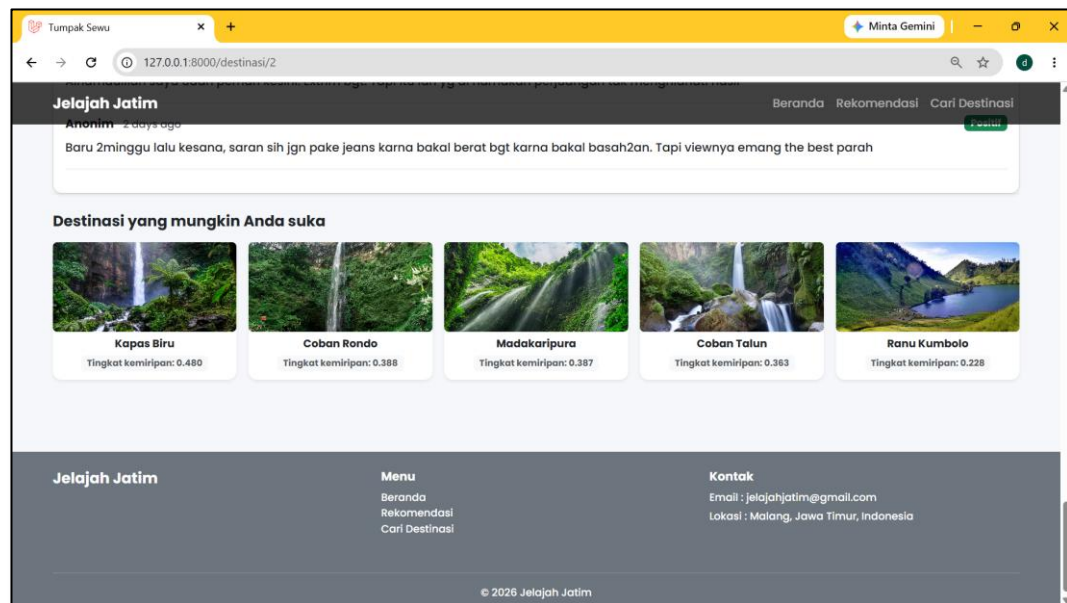
Gambar 4.9 Tampilan Halaman Rincian Destinasi

Gambar 4.9 menunjukkan *interface* halaman rincian destinasi Tumpak Sewu yang memuat foto, kategori, lokasi, dan deskripsi. Dibawah halaman rincian destinasi terdapat kesan pengunjung yang menunjukkan presentase dari ulasan yang diberikan wisatawan yang pernah berkunjung seperti pada Gambar 4.10.



Gambar 4.10 Tampilan Halaman Rincian Destinasi Kesan Pengunjung

Selain itu pada halaman ini juga menunjukkan tampilan halaman komentar wisatawan serta label sentimennya, supaya pengguna dapat melihat ulasan dari wisatawan yang pernah berkunjung. Pada bagian bawah halaman rincian destinasi dan komentar, pengguna dapat melihat lima rekomendasi destinasi wisata yang paling relevan berdasarkan tingkat kemiripan konten dengan destinasi yang dipilih, untuk memberikan alternatif pilihan sesuai preferensi pengguna.



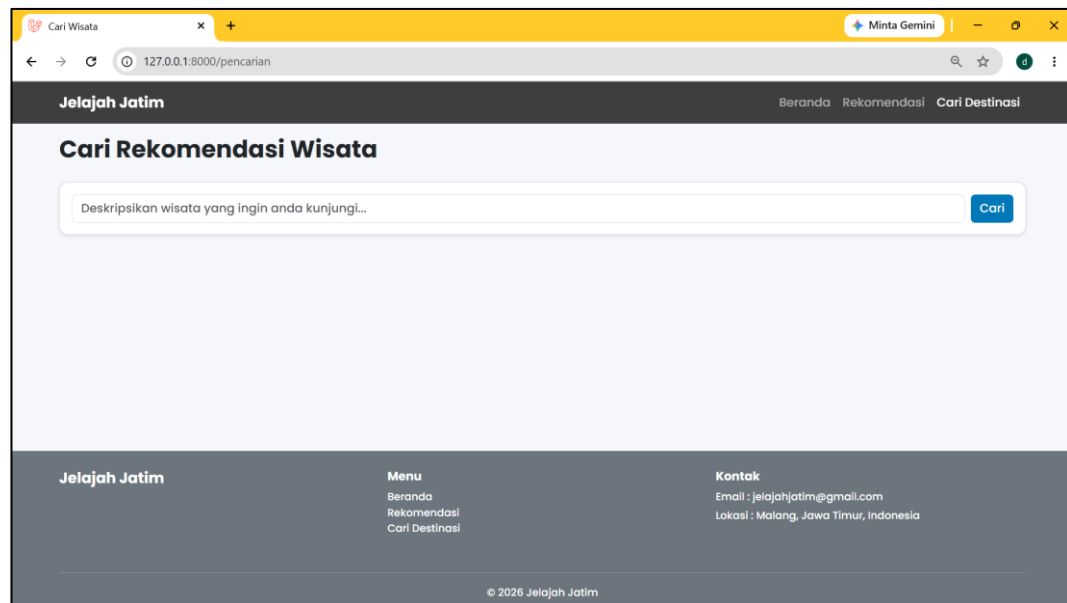
Gambar 4.11 Tampilan Rekomendasi

Gambar 4.11 menunjukkan *interface* bagian rekomendasi yang menyajikan destinasi-destinasi serupa dalam *format* kartu. Selain itu, setiap kartu destinasi juga dilengkapi dengan skor *aimilarity* untuk mengetahui tingkat kemiripan konten Tumpak Sewu dengan wisata yang direkomendasikan.

4.4.5 Halaman Pencarian Rekomendasi

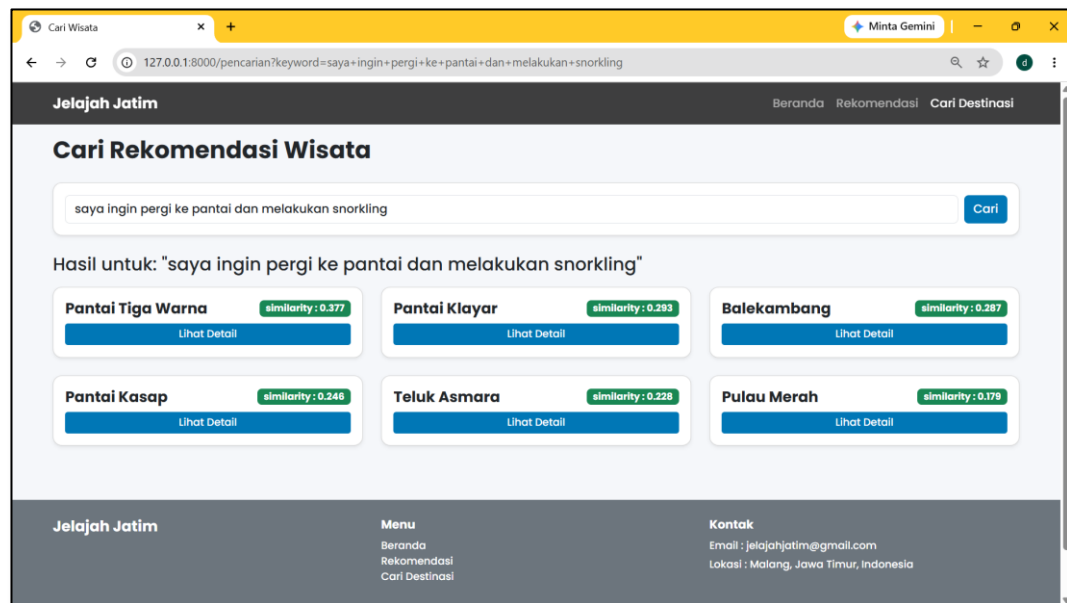
Halaman pencarian rekomendasi berfungsi sebagai antarmuka bagi pengguna untuk memperoleh rekomendasi destinasi wisata berdasarkan deskripsi yang telah diinputkan. Bagian atas halaman terdapat *form* pencarian yang terdiri atas *field input* dan tombol 'Cari'. Pengguna dapat memasukkan kata kunci atau deskripsi terkait preferensi wisata, seperti jenis wisata, suasana, atau karakteristik tertentu. Setelah input pencarian diterima dari pengguna, sistem akan memproses data tersebut menggunakan metode TF-IDF dan *Cosine Similarity* yang diimplementasikan melalui *Application Programming Interface* (API) berbasis

FastAPI. Proses ini dilakukan untuk menghasilkan nilai kemiripan (*similarity score*) antara input pengguna dengan data destinasi wisata yang tersedia. Halaman pencarian rekomendasi sebelum dilakukan pencarian ditampilkan oleh Gambar 4.12.



Gambar 4.12 Tampilan Halaman Pencarian Rekomendasi (Sebelum Pencarian)

Gambar 4.12 menampilkan tampilan awal halaman pencarian rekomendasi dengan latar dan desain yang sederhana. Halaman ini memuat field *input* dengan *placeholder* "Deskripsikan wisata yang ingin anda kunjungi..." dan tombol 'Cari' di sebelah kanan.



Gambar 4.13 Tampilan Halaman Pencarian Rekomendasi (Setelah Pencarian)

Gambar 4.13 menunjukkan tampilan halaman setelah pengguna menginputkan deskripsi "saya ingin pergi ke pantai dan melakukan snorkling". Sistem akan menampilkan lima destinasi wisata yang memiliki tingkat kemiripan tertinggi berdasarkan hasil perhitungan dalam bentuk kartu yang memuat nama destinasi, nilai *similarity*, serta tombol 'Lihat Detail' untuk mengakses informasi lebih lanjut mengenai destinasi tersebut.

4.5 Evaluasi Sistem Rekomendasi

Evaluasi terhadap sistem rekomendasi dilakukan melalui dua pendekatan. Pertama, evaluasi akurasi menggunakan metrik *Precision@K* untuk mengukur tingkat relevansi hasil rekomendasi. Kedua, evaluasi *usability* menggunakan metode PSSUQ guna mengukur tingkat kepuasan pengguna terhadap sistem yang telah dikembangkan.

4.5.1 Evaluasi *Precision@K*

Evaluasi sistem rekomendasi dilakukan menggunakan metrik *Precision@K* dengan nilai $K=5$ untuk mengukur tingkat relevansi hasil rekomendasi terhadap preferensi pengguna. Pengujian mencakup empat skenario, yaitu dua skenario untuk fitur rekomendasi berbasis pemilihan destinasi dan dua skenario untuk fitur pencarian berbasis kata kunci. Setiap skenario melibatkan lima *input* yang diuji secara bergantian. Relevansi hasil ditentukan berdasarkan penilaian lima *anotator* terhadap kesesuaian antara keluaran sistem dan *input* yang digunakan. Suatu destinasi dinyatakan relevan apabila dipilih oleh minimal tiga dari lima *anotator*, sedangkan destinasi yang dipilih oleh kurang dari tiga *anotator* dinyatakan tidak relevan.

a. Skenario 1

Skenario pertama menguji kemampuan sistem dalam merekomendasikan destinasi wisata alam pegunungan dan air terjun. Pengujian dilakukan dengan menggunakan lima destinasi sebagai *input* secara bergantian, yaitu Gunung Bromo, Tumpak Sewu, Kawah Ijen, Madakaripura, dan Ranu Kumbolo.

Tabel 4.8 Hasil Skenario 1

No	Destinasi	Nilai Similarity				
		R1	R2	R3	R4	R5
1	Gunung Bromo	0.405	0.374	0.316	0.276	0.254
2	Tumpak Sewu	0.636	0.559	0.519	0.507	0.335
3	Kawah Ijen	0.316	0.236	0.219	0.219	0.219
4	Madakaripura	0.608	0.559	0.525	0.469	0.405
5	Ranu Kumbolo	0.385	0.374	0.368	0.368	0.335

Berdasarkan Tabel 4.8, diperoleh nilai *similarity* untuk masing-masing destinasi yang menunjukkan tingkat kemiripan antar destinasi berdasarkan *input*

yang digunakan dalam sistem. Selanjutnya, pada Tabel 4.9 ditunjukkan hasil perhitungan relevansi dan nilai *precision* untuk setiap destinasi.

Tabel 4.9 Nilai *Precision* Skenario 1

No	Destinasi	Relevansi Skenario 1					<i>Precision</i>
		R1	R2	R3	R4	R5	
1	Gunung Bromo	0	1	1	0	0	2/5
2	Tumpak Sewu	1	1	1	1	0	4/5
3	Kawah Ijen	1	0	0	0	0	1/5
4	Madakaripura	1	1	1	1	0	4/5
5	Ranu Kumbolo	0	1	0	0	0	1/5
Rata-Rata							$\frac{12}{25} = 0.48$

Rata-rata *precision* pada skenario pertama mencapai 0,48, di mana destinasi Tumpak Sewu dan Madakaripura menunjukkan performa terbaik karena mayoritas *anotator* menilai rekomendasi yang diberikan relevan berdasarkan kemiripan karakteristiknya sebagai destinasi air terjun. Sebaliknya, *input* berupa kawasan pegunungan seperti Gunung Bromo, Kawah Ijen, dan Ranu Kumbolo menunjukkan nilai *precision* yang lebih rendah. Salah satu *anotator* memberikan catatan untuk *input* Gunung Bromo bahwa "area yang bisa dijangkau cukup luas". Secara teknis, rendahnya nilai presisi pada kategori ini disebabkan oleh keterbatasan variasi data destinasi pegunungan di dalam sistem. Akibat keterbatasan tersebut, sistem memunculkan destinasi di luar kawasan gunung berdasarkan urutan nilai *similarity* untuk melengkapi top lima rekomendasi. Hal ini menyebabkan destinasi pelengkap tersebut pada akhirnya dinilai tidak relevan oleh mayoritas *anotator*.

b. Skenario 2

Skenario kedua menguji kemampuan sistem dalam merekomendasikan destinasi wisata bahari. Pengujian dilakukan dengan menggunakan lima destinasi

sebagai *input*, yaitu Pantai Papuma, Balekambang, Teluk Ijo, Pasir Putih, dan Pantai Tiga Warna.

Tabel 4.10 Hasil Skenario 2

No	Destinasi	Nilai Similarity				
		R1	R2	R3	R4	R5
1	Pantai Papuma	0.406	0.392	0.385	0.369	0.361
2	Balekambang	0.545	0.530	0.516	0.504	0.495
3	Teluk Ijo	0.424	0.354	0.330	0.319	0.242
4	Pasir Putih	0.496	0.495	0.494	0.399	0.392
5	Pantai Tiga Warna	0.504	0.410	0.379	0.356	0.338

Berdasarkan Tabel 4.10, diperoleh nilai *similarity* untuk masing-masing destinasi yang menunjukkan tingkat kemiripan antar destinasi berdasarkan *input* yang digunakan dalam sistem. Tabel 4.11 menunjukkan nilai perhitungan relevansi serta nilai *precision* untuk setiap destinasi.

Tabel 4.11 Nilai *Precision* Skenario 2

No	Destinasi	Relevansi Skenario 2					<i>Precision</i>
		R1	R2	R3	R4	R5	
1	Pantai Papuma	1	1	1	1	1	5/5
2	Balekambang	1	1	1	1	1	5/5
3	Teluk Ijo	1	1	1	1	1	5/5
4	Pasir Putih	1	1	1	1	1	5/5
5	Pantai Tiga Warna	1	1	1	1	1	5/5
Rata-Rata							$\frac{25}{25} = 1$

Nilai rata-rata *precision* pada skenario kedua mencapai nilai sempurna, yakni 1,00. Seluruh *anotator* secara konsisten menyatakan bahwa setiap rekomendasi relevan dengan *input* yang diberikan. Alasan yang disampaikan pun seragam, yaitu kelima destinasi rekomendasi sama-sama menawarkan panorama pantai dan laut dengan hamparan pasir yang luas.

c. Skenario 3

Skenario ketiga menguji kemampuan fitur pencarian dalam menghasilkan destinasi yang relevan berdasarkan kata kunci yang terdiri dari dua kata. Pengujian dilakukan dengan menggunakan lima *input* kata kunci, yaitu "pantai bersih", "air terjun", "pemandangan gunung", "wisata edukasi", dan "wahana bermain".

Tabel 4.12 Hasil Skenario 3

No	Input	Nilai Similarity				
		R1	R2	R3	R4	R5
1	Pantai bersih	0.456	0.453	0.391	0.320	0.303
2	Air terjun	0.315	0.280	0.260	0.227	0.190
3	Pemandangan gunung	0.126	0.094	0.065	0.056	0.032
4	Wisata edukasi	0.127	0.086	0.041	0.038	0.034
5	Wahana bermain	0.060	0.050	0.050	0.048	0.044

Berdasarkan Tabel 4.12, diperoleh nilai *similarity* untuk masing-masing kata kunci yang menunjukkan tingkat kemiripan antara kata kunci dengan data dalam sistem. Selanjutnya, hasil perhitungan tingkat relevansi dan nilai *precision* untuk masing-masing kata kunci disajikan pada Tabel 4.13.

Tabel 4.13 Nilai *Precision* Skenario 3

No	Input	Relevansi Skenario 3					<i>Precision</i>
		R1	R2	R3	R4	R5	
1	Pantai bersih	1	1	1	1	1	5/5
2	Air terjun	1	1	1	1	1	5/5
3	Pemandangan gunung	1	1	1	0	0	3/5
4	Wisata edukasi	1	1	1	0	1	4/5
5	Wahana bermain	0	0	1	1	0	2/5
Rata-Rata							$\frac{19}{25} = 0.76$

Rata-rata *precision* pada skenario ketiga adalah 0,76. Kata kunci "pantai bersih" dan "air terjun" meraih *precision* sempurna 5/5 karena seluruh hasil rekomendasi merupakan destinasi pantai atau air terjun. Sementara itu, kata kunci "wahana bermain" memperoleh *precision* terendah sebesar 2/5 karena hanya

destinasi yang benar-benar menyediakan wahana buatan, seperti Jatim Park 2 dan BNS yang dinilai relevan oleh mayoritas *anotator*, sedangkan destinasi lainnya dianggap tidak sesuai karena tidak memiliki fasilitas tersebut.

d. Skenario 4

Skenario keempat menguji kemampuan fitur pencarian dalam menghasilkan destinasi yang relevan berdasarkan kata kunci yang terdiri dari 5 kata. Pengujian dilakukan dengan menggunakan lima kata kunci sebagai *input*, yaitu "pantai yang tenang dan snorkeling", "jalur ekstrim dan pemandangan indah", "pemandangan gunung indah dengan kabut", "wisata edukasi cocok anak keluarga", dan "tempat wisata banyak wahana seru".

Tabel 4.14 Hasil Skenario 4

No	Input	Nilai Similarity				
		R1	R2	R3	R4	R5
1	Pantai yang tenang dan snorkeling	0.337	0.264	0.247	0.225	0.196
2	Jalur ekstrim dan pemandangan indah	0.059	0.057	0.052	0.031	0.030
3	Pemandangan gunung indah dengan kabut	0.137	0.083	0.079	0.052	0.041
4	Wisata edukasi cocok anak keluarga	0.151	0.106	0.084	0.063	0.063
5	Tempat wisata banyak wahana seru	0.355	0.262	0.087	0.086	0.060

Tabel 4.14 menyajikan nilai *similarity* untuk masing-masing kata kunci yang menunjukkan tingkat kemiripan antara kata kunci naratif dengan destinasi wisata dalam sistem. Hasil perhitungan relevansi beserta nilai *precision* untuk setiap *input* disajikan pada Tabel 4.15.

Tabel 4.15 Nilai *Precision* Skenario 4

No	Input	Relevansi Skenario 4					<i>Precision</i>
		R1	R2	R3	R4	R5	
1	Pantai yang tenang dan snorkeling	1	0	0	0	1	2/5
2	Jalur ekstrim dan pemandangan indah	0	1	0	0	0	1/5
3	Pemandangan gunung indah dengan kabut	1	1	1	0	0	3/5
4	Wisata edukasi cocok anak keluarga	1	1	0	0	1	3/5
5	Tempat wisata banyak wahana seru	1	1	1	1	0	4/5
Rata-Rata							$\frac{13}{25} = 0.52$

Rata-rata *precision* pada skenario keempat adalah 0,52. Frasa "tempat wisata banyak wahana seru" meraih *precision* tertinggi sebesar 4/5 karena mayoritas *anotator* menilai destinasi yang direkomendasikan relevan, yakni menawarkan beragam wahana edukatif maupun rekreatif. Sebaliknya, frasa "jalur ekstrim dan pemandangan indah" memperoleh *precision* terendah sebesar 1/5 karena sebagian besar rekomendasi dinilai tidak relevan karena berupa destinasi yang dapat diakses dengan kendaraan roda empat, yang mana kurang merepresentasikan frasa 'jalur ekstrim' pada kata kunci *input*.

Tabel 4.16 Nilai Rata-rata *Precision*

No	Skenario	Jumlah Relevan	<i>Precision</i>
1	Skenario 1	12	0.48
2	Skenario 2	25	1.00
3	Skenario 3	19	0.76
4	Skenario 4	13	0.52
Rata-rata			$\frac{0.48 + 1.00 + 0.76 + 0.52}{4} = 0.69$

Berdasarkan Tabel 4.16, rata-rata *precision* dari keempat skenario adalah 0,69. Fitur rekomendasi pada skenario wisata bahari mencatat nilai tertinggi sebesar 1,00, sedangkan skenario wisata alam pegunungan dan air terjun memperoleh nilai terendah sebesar 0,48. Menurut penilaian *anotator*, rendahnya nilai pada kategori wisata alam disebabkan oleh keberagaman karakteristik yang dicakup, mulai dari

gunung berapi, air terjun, kawah, hingga danau pegunungan, sehingga rekomendasi dari satu sub-kategori, misalnya air terjun, kerap dinilai tidak relevan terhadap *input* dari sub-kategori berbeda seperti gunung. Kondisi ini berbeda dengan wisata bahari yang karakteristiknya dinilai lebih homogen oleh para *anotator*.

4.5.2 Evaluasi *Usability* Menggunakan PSSUQ

Evaluasi *usability* sistem dilakukan menerapkan pendekatan PSSUQ dalam mengukur tingkat kepuasan pengguna terhadap sistem berdasarkan pengalaman mereka setelah menggunakan sistem secara langsung. Pengujian ini melibatkan sebanyak 52 responden yang telah mencoba sistem yang dikembangkan, di mana responden diminta mengisi kuesioner PSSUQ setelah berinteraksi langsung dengan sistem.

Kuesioner yang digunakan terdiri dari 16 pertanyaan yang terbagi ke dalam empat aspek utama, yaitu *System Usefulness* (SysUse), *Information Quality* (InfoQual), *Interface Quality* (InterQual), dan *Overall*, dengan skala penilaian *likert* dengan ketentuan bahwa skor yang lebih rendah mengindikasikan tingkat *usability* sistem yang lebih baik. Hasil pengujian kemudian dibandingkan dengan norma skala PSSUQ pada Tabel 4.17.

Tabel 4.17 Norma Skala PSSUQ

Skala	Aturan Penilaian Skala	Batas Bawah	Mean	Batas Atas
SysUse	Rata Pertanyaan 1-6	2.57	2.80	3.62
InfoQual	Rata Pertanyaan 7-12	2.79	3.02	3.24
InterQual	Rata Pertanyaan 13-15	2.28	2.49	2.71
<i>Overall</i>	Rata Pertanyaan 1-16	2.62	2.82	3.02

Berdasarkan tabel norma tersebut, suatu sistem dikatakan memiliki tingkat *usability* yang baik apabila nilai rata-rata tiap skala yang diperoleh berada di bawah norma skala nilai mean, dan semakin baik jika berada di bawah batas bawah.

4.5.2.1 *System Usefulness (SysUse)*

Aspek *System Usefulness* dapat mengukur sejauh mana sistem mampu mendukung pengguna dalam menyelesaikan tugas secara efektif dan efisien. Penilaian terhadap aspek ini dihitung berdasarkan nilai rata-rata dari butir pertanyaan P1 sampai dengan P6.

Tabel 4.18 Metode PSSUQ Aspek *System Usefulness*

Responden	P1	P2	P3	P4	P5	P6	SysUse
R1	1	1	1	2	1	1	1.167
R2	2	1	1	2	3	2	1.833
R3	2	2	2	3	3	2	2.333
R4	1	3	3	4	1	2	2.333
R5	1	1	1	3	1	1	1.333
...	...	...	...	...	...	...	...
R50	4	3	3	3	3	3	3.167
R51	2	1	2	2	2	3	2
R52	1	1	1	1	1	1	1
Rata-Rata							2.02

Menurut Tabel 4.18, nilai rata-rata *System Usefulness* sebesar 2,02, lebih rendah dari ambang batas norma PSSUQ yaitu 2,57. Hasil evaluasi tersebut menunjukkan bahwa performa sistem dinilai sangat berguna karena berhasil mendukung pengguna dalam menyelesaikan pekerjaan dengan lebih efektif dan efisien.

4.5.2.2 *Information Quality (InfoQual)*

Aspek *Information Quality* dirancang untuk mengevaluasi mutu informasi yang disediakan oleh sistem, yang mencakup parameter kejelasan, kelengkapan,

serta tingkat keterpahaman data bagi pengguna. Pengukuran pada dimensi ini dihitung berdasarkan nilai rata-rata dari butir pertanyaan P7 sampai P12.

Tabel 4.19 Metode PSSUQ Aspek *Information Quality*

Responden	P7	P8	P9	P10	P11	P12	InfoQual
R1	2	1	1	1	3	1	1.5
R2	1	3	2	2	3	1	2
R3	2	2	2	2	3	2	2.167
R4	1	2	1	2	2	2	1.667
R5	1	1	1	1	3	1	1.333
...	...	...	...	...	...	...	...
R50	3	3	3	3	3	3	3
R51	4	3	2	2	1	3	2.5
R52	1	1	1	1	1	1	1
Rata-Rata							2.14

Menurut Tabel 4.19, skor rata-rata *Information Quality* sebesar 2,14, lebih rendah dari ambang batas norma PSSUQ yaitu 2,79. Capaian ini mengindikasikan bahwa pengguna menganggap informasi dalam sistem sudah tersampaikan secara gamblang, menyeluruh, dan mudah dipahami.

4.5.2.3 *Interface Quality (InterQual)*

Aspek *Interface Quality* digunakan untuk mengevaluasi kualitas tampilan antarmuka sistem, termasuk kenyamanan, konsistensi, dan kemudahan dalam berinteraksi dengan sistem. Penilaian ini didasarkan pada rata-rata dari pertanyaan P13 hingga P15.

Tabel 4.20 Metode PSSUQ Aspek *Interface Quality*

Responden	P13	P14	P15	InterQual
R1	3	2	1	2
R2	1	2	2	1.667
R3	2	2	2	2
R4	1	1	2	1.333
R5	1	1	1	1
...	...	...	...	...
R50	3	3	3	3
R51	4	4	3	3.667
R52	1	1	1	1
Rata-Rata				2.05

Menurut data pada Tabel 4.20, skor rata-rata *Interface Quality* sebesar 2,05, lebih rendah dari ambang batas norma PSSUQ yaitu 2,28. Hasil ini menjadi bukti bahwa tampilan visual sistem dinilai ramah pengguna, nyaman, serta mudah untuk digunakan.

4.5.2.4 Overall

Aspek *Overall* merupakan penilaian keseluruhan terhadap *usability* sistem yang dihitung berdasarkan rata-rata seluruh pertanyaan dari P1 hingga P16.

Tabel 4.21 Metode PSSUQ Aspek *Overall*

Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	Overall
R1	1	1	1	2	1	1	2	1	1	1	3	1	3	2	1	1	1.438
R2	2	1	1	2	3	2	1	3	2	2	3	1	1	2	2	2	1.875
R3	2	2	2	3	3	2	2	2	2	2	3	2	2	2	2	3	2.25
R4	1	3	3	4	1	2	1	2	1	2	2	2	1	1	2	1	1.813
R5	1	1	1	3	1	1	1	1	1	1	3	1	1	1	1	1	1.25
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
R50	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3.063
R51	2	1	2	2	2	3	4	3	2	2	1	3	4	4	3	2	2.5
R52	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Rata-Rata																	2.07

Dari analisis data yang dilakukan, diperoleh skor rata-rata *Overall* sebesar 2,07. Karena angka ini lebih rendah dari batas norma PSSUQ (2,62), dapat disimpulkan bahwa sistem memiliki kualitas *usability* yang unggul. Pengguna merasa terbantu oleh kemudahan akses sistem, kualitas informasi yang disajikan, hingga kenyamanan antarmukanya. Hasil ini membuktikan bahwa sistem telah memenuhi standar fungsi secara menyeluruh dan sukses mencapai tingkat kepuasan pengguna yang optimal.

4.6 Integrasi Islam

Integrasi nilai-nilai Islam dalam penelitian ini bertujuan untuk memberikan landasan etis dan spiritual dalam pengembangan sistem rekomendasi wisata. Islam tidak hanya mengatur aspek ibadah, tetapi juga mengaitkan dimensi spiritualitas kepada Allah SWT, dimensi sosial terhadap sesama, serta kepedulian pada alam semesta. Oleh karena itu, penelitian ini diharapkan menghasilkan sistem yang bermanfaat secara teknis sekaligus selaras dengan prinsip-prinsip keislaman. Integrasi tersebut dijabarkan melalui dua dimensi utama, yaitu *Mu'ammalah Ma'allah* dan *Mu'ammalah Ma'annas*.

4.6.1 *Mu'ammalah Ma'allah* (Hubungan Manusia dengan Allah)

Perspektif Islam memandang bahwa hubungan makhluk dengan Penciptanya (*Mu'amalah Ma'allah*) dibuktikan lewat iman dan takwa. Bentuk nyatanya adalah dengan mengagumi sekaligus merenungkan alam ciptaan Allah, termasuk menjadikannya tujuan utama saat melakukan perjalanan wisata.

﴿قُلْ سِيرُوا فِي الْأَرْضِ فَانظُرُوا كَيْفَ بَدَأَ الْخَلْقَ ثُمَّ اللَّهُ يُنشِئُ النَّشْأَةَ الْآخِرَةَ إِنَّ اللَّهَ عَلَىٰ كُلِّ شَيْءٍ قَدِيرٌ ۝ ٢٠﴾

“Katakanlah, berjalanlah di (muka) bumi, lalu perhatikanlah bagaimana Allah memulai penciptaan (semua makhluk). Kemudian, Allah membuat kejadian yang akhir (setelah mati di akhirat kelak). Sesungguhnya Allah Mahakuasa atas segala sesuatu.” (QS. Al-Ankabut: 20)

Surah Al-Ankabut ayat 20 mengajak manusia untuk menjelajahi bumi dan melihat langsung bagaimana Allah menciptakan segala sesuatu. Perjalanan ini bukan sekadar untuk jalan-jalan atau bersenang-senang biasa, tetapi juga untuk belajar dan mengambil hikmah dari alam semesta. Dalam tradisi Pendidikan islam,

aktivitas ini dikenal sebagai *rihlah*, yaitu perjalanan yang bertujuan memperluas wawasan dan mempertebal ketauhidan (Wahidi, 2016).

Dalam penelitian ini, sistem rekomendasi yang dibuat berfungsi untuk memudahkan pengguna mencari tempat wisata alam yang sesuai dengan keinginan mereka. Dengan memberikan rekomendasi destinasi yang tepat, sistem ini secara tidak langsung membantu memfasilitasi pengguna untuk melakukan perjalanan yang bermakna. Sistem membuat perencanaan wisata menjadi lebih mudah, sehingga pengguna dapat lebih fokus menikmati keindahan alam dan mensyukuri kebesaran ciptaan Allah SWT sesuai anjuran ayat tersebut.

4.6.2 *Mu'ammalah Ma'annas* (Hubungan Manusia dengan Sesama Manusia)

Dalam Islam, hubungan manusia dengan sesama (*Mu'ammalah Ma'annas*) menekankan pentingnya kejujuran, tanggung jawab, serta kehati-hatian dalam menyampaikan dan menerima informasi. Setiap informasi harus memiliki dasar yang jelas dan dapat dipertanggungjawabkan.

﴿وَلَا تَقْفُ مَا لَيْسَ لَكَ بِهِ عِلْمٌ إِنَّ السَّمْعَ وَالْبَصَرَ وَالْفُؤَادَ كُلُّ أُولَٰئِكَ كَانَ عَنْهُ مَسْئُولًا ۚ ٣٦﴾

“Janganlah engkau mengikuti sesuatu yang tidak kauketahui. Sesungguhnya pendengaran, penglihatan, dan hati nurani, semua itu akan diminta pertanggungjawabannya.” (QS. Al-Isra’: 36)

Surah Al-Isra ayat 36 menegaskan larangan mengikuti atau menyampaikan sesuatu tanpa pengetahuan yang jelas. Dalam tafsir Ibnu Katsir (2004) dijelaskan bahwa manusia tidak diperkenankan berpendapat atau mengambil keputusan tanpa dasar ilmu yang valid. Oleh karena itu, setiap informasi perlu diverifikasi sebelum digunakan sebagai dasar pengambilan keputusan.

Prinsip kehati-hatian ini sangat sejalan dengan alasan mengapa sistem rekomendasi ini dibuat. Saat mencari tempat wisata di internet, pengguna berisiko terjebak oleh gambar atau promosi wisata yang tidak sesuai dengan kondisi aslinya. Untuk mencegah hal tersebut, sistem ini menampilkan komentar atau ulasan langsung dari wisatawan yang pernah berkunjung. Fitur ini menjadi ruang bagi sesama pengguna untuk saling memberikan kesaksian yang jujur. Dengan membaca pengalaman nyata dari orang lain, pengguna baru bisa mendapatkan gambaran lokasi wisata secara apa adanya. Hal ini merupakan bentuk tolong-menolong dalam membagikan informasi yang benar antar sesama manusia.

Selain itu, sistem juga menerapkan metode TF-IDF dan *Cosine Similarity* untuk menyaring dan menghitung kecocokan data ulasan tersebut secara matematis. Melalui perpaduan ulasan asli pengguna dan perhitungan algoritma yang akurat, sistem ini bertindak sebagai alat penyaring informasi yang jujur. Sistem memastikan pengguna tidak "mengikuti sesuatu tanpa pengetahuan", sehingga mereka terhindar dari kekecewaan akibat informasi wisata yang menyesatkan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari seluruh hasil penelitian dan pembahasan di atas, diperoleh beberapa kesimpulan penting sebagai berikut.

1. Penelitian ini telah berhasil membangun sistem rekomendasi destinasi wisata di Jawa Timur menggunakan metode *Content-Based Filtering*. Lewat pemanfaatan representasi teks berbasis TF-IDF serta algoritma *Cosine Similarity* untuk mengukur kemiripan objek wisata, sistem ini terbukti mampu menyajikan pilihan tempat liburan yang relevan dan sesuai dengan keinginan pengguna.
2. Pengembangan fitur pencarian destinasi wisata berhasil mengintegrasikan ulasan wisatawan untuk memperkaya konten dalam proses perhitungan TF-IDF dan *Cosine Similarity*, sehingga hasil pencarian tidak hanya bergantung pada deskripsi konten destinasi, tetapi juga mencerminkan pengalaman nyata wisatawan.
3. Berdasarkan evaluasi metrik *Precision@K* ($K=5$), sistem rekomendasi menghasilkan rata-rata *precision* keseluruhan sebesar 0,69. Pada fitur rekomendasi, wisata bahari meraih hasil sempurna (1,00), sedangkan wisata alam pegunungan dan air terjun lebih rendah (0,48) akibat karakteristik data yang terlalu beragam. Sementara itu pada fitur pencarian, kata kunci pendek menghasilkan *precision* 0,76, lebih baik daripada kata kunci panjang yang hanya mencapai 0,52. Secara keseluruhan, hasil ini memaparkan bahwa

sistem mampu memberikan rekomendasi yang sesuai dan relevan, meskipun performanya dipengaruhi oleh homogenitas kategori destinasi dan panjang kata kunci input.

4. Evaluasi *usability* dengan PSSUQ terhadap 52 responden menghasilkan skor rata-rata 2,02 untuk *System Usefulness*, 2,14 untuk *Information Quality*, 2,05 untuk *Interface Quality*, dan 2,07 untuk *Overall*. Seluruh nilai tersebut berhasil berada di bawah batas bawah norma skala PSSUQ. Hal ini membuktikan bahwa sistem memiliki tingkat *usability* yang sangat baik serta mampu memberikan kepuasan pengalaman pengguna yang optimal.

5.2 Saran

Melihat hasil yang dicapai dalam penelitian ini, beberapa saran berikut dapat dijadikan acuan untuk proses pengembangan sistem selanjutnya. Cakupan data destinasi wisata perlu diperluas, terutama untuk destinasi yang belum banyak dikenal, serta menyajikan informasi destinasi secara lebih lengkap, mencakup tarif masuk, jam buka-tutup tempat wisata, serta sarana prasarana yang ada di sana. Fitur pencarian dapat dikembangkan lebih lanjut dengan memperbanyak data ulasan wisatawan agar representasi konten destinasi semakin kaya dan hasil pencarian semakin relevan.

DAFTAR PUSTAKA

- Astuti, F. D., & Andriyani, W. (2025). Pengembangan Sistem Rekomendasi Pembimbing Tugas Akhir Menggunakan Teknik Content-Based Filtering. *JIKO (Jurnal Informatika Dan Komputer)*, 9(2), 474. <https://doi.org/10.26798/jiko.v9i2.1599>
- Eddyono, F. (2021). *Pengelolaan Destinasi Pariwisata* (1st ed.). Uwais Inspirasi Indonesia.
- Hamka. (1990). *Tafsir Al-Azhar Jilid 2*. (Original work published Pustaka Nasional PTE LTD Singapura)
- Howay, S., & Rianto, R. (2021). Sistem Rekomendasi Jurusan pada Sekolah Menengah Kejuruan (SMK) dengan Algoritma K-Means. *Syntax Idea*, 3(10), 2215–2222. <https://doi.org/10.46799/syntax-idea.v3i10.1443>
- Ilhamsyah, I., Rahmayuda, S., Midyanti, D. M., & Martha, S. (2024). Pemodelan Sistem Rekomendasi Restoran berdasarkan Preferensi Pengguna dengan Pendekatan Content-Based Filtering. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 10(1), 154. <https://doi.org/10.26418/jp.v10i1.74008>
- Indriani, M., Ahsyar, T. K., Syaifullah, S., & Fronita, M. (2024). Evaluasi Usability Aplikasi Streaming Film Menggunakan Metode Post Study System Usability Questionnaire (PSSUQ). *Jurnal Algoritma*, 21(2), 221–231. <https://doi.org/10.33364/algoritma/v.21-2.1660>
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). *Recommender Systems An Introduction*. Cambridge University Press.
- Katsir, I. I. (2004). *Tafsir Ibnu Katsir Juz 15*. Sinar Baru Algensindo.
- Katsir, I. I. (2021a). *Terjemah Tafsir Ibnu Katsir Al-Hujurat* (14th ed., Vol. 8). Jakarta : Pustaka Imam Asy-Syafii.
- Katsir, I. I. (2021b). *Terjemah Tafsir Ibnu Katsir Al-Mulk* (14th ed., Vol. 10). Jakarta : Pustaka Imam Asy-Syafii.
- Maulana, A., Rohman, A. N., & Pristyanto, Y. (2026). *Sistem Rekomendasi Destinasi Tempat Wisata Provinsi Aceh menggunakan Metode Content based Filtering*. 15.
- Muahidin, Z., & Azkiyah, E. (2024). Penerapan Algoritma Collaborative Filtering pada Sistem Informasi Penyewaan Jasa Tour Guide Berbasis Web di Wisata Benang Kelambu. *Journal Computer and Technology*, 2(2), 83–88. <https://doi.org/10.69916/comtechno.v2i2.148>
- Nastiti, P. (2019). Penerapan Metode Content Based Filtering Dalam Implementasi Sistem Rekomendasi Tanaman Pangan. *Teknika*, 8(1), 1–10. <https://doi.org/10.34148/teknika.v8i1.139>
- Oyadila, S., Abdullah, D., & Razi, A. (2025). Implementasi Content-Based Filtering Dengan TF-IDF Dan Cosine Similarity Untuk Sistem Rekomendasi Destinasi Wisata Di Aceh Tengah. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 10(2), 1329–1339. <https://doi.org/10.36341/rabit.v10i2.6532>

- Palelleng, S., Adha, M. S., Pasinggi, E. S., & Muhammad, M. (2025). Pengembangan Aplikasi Edukasi Wisata Berbasis Smart Tourism Dengan Sistem Rekomendasi Content-Based Filtering. *VOX EDUKASI: Jurnal Ilmiah Ilmu Pendidikan*, 16(2), 513–523. <https://doi.org/10.31932/ve.v16i2.5698>
- Pariwisata, K. (2025a, September 1). *Statistik Kunjungan Wisatawan Mancanegara Bulan Juli 2025*. Kementerian Wisata Republik Indonesia. <https://kemenpar.go.id/direktori-statistik/statistik-kunjungan-wisatawan-mancanegara-bulan-juli-2025>
- Pariwisata, K. (2025b, September 3). *Statistik Perjalanan Wisatawan Nusantara Bulan Juli Tahun 2025*. Kementerian Wisata Republik Indonesia. <https://kemenpar.go.id/direktori-statistik/statistik-perjalanan-wisatawan-nusantara-bulan-juli-tahun-2025>
- Pradana, D. S., Prajoko, P., & Hartawan, G. P. (2022). Perbandingan Algoritma Content-Based Filtering dan Collaborative Filtering dalam Rekomendasi Kegiatan Ekstrakurikuler Siswa. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 151. <https://doi.org/10.35889/progresif.v18i2.854>
- Pratiwi, D., Asrianda, A., & Rosnita, L. (2025). Penerapan Metode Content-Based Filtering dalam Sistem Rekomendasi Objek Wisata di Aceh Tamiang. *Jurnal Ilmu Komputer Dan Informatika*, 4(2), 85–96. <https://doi.org/10.54082/jiki.169>
- Putri, N. I., Rustiyana, Herdiana, Y., & Zen Munawar. (2021). Sistem Rekomendasi Hibrid Pemilihan Mobil Berdasarkan Profil Pengguna dan Profil Barang. *TEMATIK*, 8(1), 56–68. <https://doi.org/10.38204/tematik.v8i1.566>
- Salsabilla, C., & Utomo, D. W. (2025). Sistem Rekomendasi Wisata Kabupaten Pekalongan dengan Content Based Filtering. *SISTEMASI*, 14(1), 262. <https://doi.org/10.32520/stmsi.v14i1.4839>
- Santoti, J. V., Jocelyn, J., & Irsyad, H. (2025). Implementasi Term Frequency—Inverse Document Frequency dan Cosine Similarity untuk Analisis Kemiripan Deskripsi Produk Halal. *Jurnal Software Engineering and Computational Intelligence*, 3(01), 44–52. <https://doi.org/10.36982/jseci.v3i01.5421>
- Suryadi, I., Gutama, D. H., Danianti, D., & Yazid, A. S. (2025). *Sistem Rekomendasi Tempat Wisata Alam di Gunung Kidul Menggunakan Algoritma Content Based Filtering Berbasis Web*. 9(2). <https://doi.org/10.36040/jati.v9i2.13344>
- Suwandy, R., Marpaung, S. H., & Caroline, C. (2022). *Evaluasi Pengalaman Pengguna dengan Menggunakan Post Study System Usability Questionnaire (PSSUQ) Perpustakaan Digital Universitas Mikroskil*. 6. <https://doi.org/https://doi.org/10.31004/jptam.v6i1.3521>
- Wahidi, R. (2016). *Tafsir Ayat-Ayat Tarbawi* (Vol. 1). (Original work published Trussmedia Grafika)
- Wulandari, F., & Husni Hoer, U. (2025). Konsep Belajar Menurut Pandangan Imam Al-Ghazali Perspektif Hadist Tarbawi. *Epistemic: Jurnal Ilmiah Pendidikan*, 4(1), 56–71. <https://doi.org/10.70287/epistemic.v4i1.384>
- Yaya Suharya, Herdiana, Y., Indah Putri, N., & Munawar, Z. (2021). Sistem Rekomendasi untuk Toko Online Kecil dan Menengah. *TEMATIK*, 8(2), 176–185. <https://doi.org/10.38204/tematik.v8i2.683>

Yulfihani, I., & Zakariyah, M. (2024). Optimization of Tourism Destination Recommendations in Batang Regency Using Content-Based Filtering. *Journal of Applied Informatics and Computing*, 8(2), 499–508. <https://doi.org/10.30871/jaic.v8i2.8618>

LAMPIRAN

Lampiran I. Seluruh Data Destinasi Wisata

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
1	Gunung Bromo	Probolinggo	Wisata Alam Pegunungan Dan Air Terjun	Gunung Bromo merupakan gunung berapi aktif di Jawa Timur yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Destinasi ini dikenal dengan lanskap pegunungan vulkanik, lautan pasir luas, serta kawah aktif yang menjadi daya tarik utama wisata alam. Beberapa lokasi populer di kawasan ini antara lain Kawah Bromo, Pasir Berbisik, dan Bukit Teletubbies.
2	Tumpak Sewu	Lumajang	Wisata Alam Pegunungan Dan Air Terjun	Air Terjun Tumpak Sewu merupakan destinasi wisata alam yang terletak di wilayah perbatasan Kabupaten Lumajang dan Kabupaten Malang. Air terjun ini dikenal memiliki aliran air yang menyerupai tirai dengan latar tebing tinggi dan lingkungan alam pegunungan. Kawasan ini menjadi daya tarik wisata alam karena keindahan lanskap dan karakter air terjunnya yang unik.
3	Jatim Park 1	Batu	Wisata Buatan Dan Rekreasi	Jawa Timur Park 1 merupakan destinasi wisata buatan di Kota Batu yang mengusung konsep taman hiburan berbasis edukasi dan rekreasi. Kawasan ini dikenal sebagai taman sains dan permainan yang menyajikan berbagai wahana edukatif serta hiburan keluarga. Karakter destinasi ini memadukan unsur pembelajaran dan rekreasi dalam satu kawasan wisata.
4	Pantai Papuma	Jember	Wisata Bahari	Pantai Papuma atau Pantai Tanjung Papuma merupakan destinasi wisata bahari di Kabupaten Jember yang memiliki pasir putih, batu karang besar, dan kawasan hutan lindung di sekitarnya. Pantai ini dikenal sebagai lokasi wisata alam dengan panorama laut selatan serta pemandangan matahari terbit dan terbenam. Berbagai aktivitas wisata alam dapat dilakukan di kawasan pantai ini dengan dukungan fasilitas penunjang wisata.
5	Baluran	Situbondo	Wisata Hutan Dan Konservasi Alam	Taman Nasional Baluran merupakan kawasan konservasi alam di Jawa Timur yang dikenal dengan bentang alam savana luas dengan latar Gunung Baluran. Kawasan ini memiliki ekosistem savana, hutan, dan pesisir yang menjadi habitat berbagai satwa liar seperti banteng, rusa, dan berbagai

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				jenis burung. Daya tarik utama kawasan ini adalah Savana Bekol serta keberagaman lanskap alam yang khas.
6	Kawah Ijen	Banyuwangi	Wisata Alam Pegunungan Dan Air Terjun	Kawah Ijen merupakan destinasi wisata alam di Jawa Timur yang terkenal dengan fenomena api biru atau blue fire. Kawasan ini memiliki danau kawah dengan kandungan belerang tinggi serta dikelilingi tebing kawah yang membentuk lanskap vulkanik. Aktivitas penambangan belerang dan jalur pendakian menjadi bagian dari karakter kawasan wisata alam ini.
7	Balekambang	Malang	Wisata Bahari	Pantai Balekambang merupakan destinasi wisata bahari di Kabupaten Malang yang berada di kawasan pesisir selatan Jawa Timur. Pantai ini dikenal dengan garis pantai berkarang serta keberadaan pulau-pulau kecil di lepas pantai. Karakter alam pesisir dan lanskap pantai menjadi daya tarik utama kawasan wisata ini.
8	Teluk Asmara	Malang	Wisata Bahari	Pantai Teluk Asmara merupakan destinasi wisata bahari yang memiliki pasir putih dan gugusan pulau-pulau kecil di sekitarnya. Pantai ini dikelilingi bukit kapur yang membentuk teluk dengan perairan yang relatif tenang. Karakter kawasan pantai ini menjadikannya sebagai lokasi wisata alam dengan suasana pantai yang alami.
9	Pulau Merah	Banyuwangi	Wisata Bahari	Pulau Merah merupakan destinasi wisata pantai di Kabupaten Banyuwangi yang ditandai dengan keberadaan bukit kecil berwarna kemerahan di dekat garis pantai. Pantai ini memiliki bentang alam pesisir terbuka dengan panorama laut selatan. Karakter pantai dan lanskap alamnya menjadi ciri khas destinasi wisata bahari ini.
10	Coban Talun	Batu	Wisata Alam Pegunungan Dan Air Terjun	Coban Talun merupakan destinasi wisata air terjun di Kota Batu yang berada di kawasan pegunungan dengan lingkungan alam yang sejuk. Air terjun ini memiliki aliran air jernih dan dikelilingi kawasan hutan pegunungan. Karakter alamnya menjadikan Coban Talun sebagai destinasi wisata alam pegunungan.
11	Pasir Putih	Situbondo	Wisata Bahari	Pantai Pasir Putih Situbondo merupakan destinasi wisata bahari di Kabupaten Situbondo yang dikenal

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				dengan pantai berpasir putih, garis pantai yang landai, dan ombak yang relatif tenang. Kawasan pantai ini memiliki perairan dangkal yang sesuai untuk aktivitas wisata pantai dan rekreasi alam. Karakter pantai yang terbuka menjadikannya salah satu destinasi wisata pesisir di wilayah utara Jawa Timur.
12	Alas Purwo	Banyuwangi	Wisata Hutan Dan Konservasi Alam	Taman Nasional Alas Purwo merupakan kawasan konservasi alam di Kabupaten Banyuwangi yang memiliki ekosistem hutan hujan tropis, savana, mangrove, dan pesisir pantai. Kawasan ini menjadi habitat berbagai flora dan fauna liar serta memiliki nilai konservasi dan pelestarian lingkungan. Karakter kawasan yang masih alami menjadikannya sebagai destinasi wisata alam berbasis konservasi.
13	Selecta	Batu	Wisata Buatan Dan Rekreasi	Selecta merupakan destinasi wisata buatan di Kota Batu yang dikenal dengan kawasan taman bunga dan pemandian rekreasi. Kawasan ini berada di lingkungan pegunungan dengan udara sejuk dan panorama alam sekitar. Karakter destinasi ini memadukan wisata taman dan rekreasi keluarga dalam satu kawasan.
14	Taman Safari Prigen	Pasuruan	Wisata Hutan Dan Konservasi Alam	Taman Safari Prigen merupakan destinasi wisata berbasis konservasi satwa di Jawa Timur yang mengusung konsep safari alam terbuka. Kawasan ini menampilkan berbagai jenis satwa dalam lingkungan buatan yang menyerupai habitat aslinya. Karakter destinasi ini memadukan wisata alam, konservasi, dan edukasi satwa.
15	Pantai Klayar	Pacitan	Wisata Bahari	Pantai Klayar merupakan destinasi wisata bahari di Kabupaten Pacitan yang dikenal dengan bentang alam pantai berkarang dan <i>formasi</i> batuan alami. Pantai ini memiliki fenomena celah batu karang yang dapat menyemburkan air laut saat ombak besar datang. Karakter alam pesisir yang unik menjadi daya tarik utama kawasan pantai ini.
16	Ranu Kumbolo	Lumajang	Wisata Alam Pegunungan Dan Air Terjun	Ranu Kumbolo merupakan danau air tawar yang berada di kawasan Taman Nasional Bromo Tengger Semeru. Danau ini terletak di kaki Gunung Semeru dan dikelilingi lanskap

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				perbukitan serta hutan pegunungan. Kondisi alam yang masih terjaga menjadikan Ranu Kumbolo sebagai bagian dari wisata alam pegunungan.
17	Jatim Park 2	Batu	Wisata Buatan Dan Rekreasi	Jawa Timur Park 2 merupakan destinasi wisata buatan di Kota Batu yang mengusung konsep taman hiburan berbasis konservasi dan edukasi. Kawasan ini terdiri dari beberapa area utama seperti kebun binatang tematik dan museum satwa yang menampilkan berbagai koleksi fauna. Karakter destinasi ini menekankan perpaduan antara wisata rekreasi dan edukasi lingkungan.
18	Batu Night Sprctacular	Batu	Wisata Buatan Dan Rekreasi	Batu Night Spectacular merupakan destinasi wisata buatan di Kota Batu yang mengusung konsep hiburan malam. Kawasan ini menggabungkan berbagai wahana permainan dan atraksi hiburan dalam satu area wisata. Karakter destinasi ini menonjolkan suasana rekreasi malam sebagai daya tarik utama.
19	Museum Angkut	Batu	Wisata Buatan Dan Rekreasi	Museum Angkut merupakan destinasi wisata edukasi di Jawa Timur yang menampilkan koleksi berbagai moda transportasi dari tradisional hingga modern. Kawasan museum ini dikemas dengan konsep tematik yang menggabungkan koleksi kendaraan dan replika lingkungan dari berbagai negara. Museum ini menjadi destinasi wisata buatan yang mengedepankan aspek edukasi dan rekreasi.
20	Pantai 3 Warna	Malang	Wisata Bahari	Pantai Tiga Warna merupakan destinasi wisata bahari di Kabupaten Malang yang berada dalam kawasan konservasi lingkungan. Pantai ini dikenal dengan gradasi warna air laut yang berbeda serta kondisi ekosistem pesisir yang masih terjaga. Pengelolaan kawasan pantai ini menekankan pada wisata alam berbasis konservasi.
21	Madakaripura	Probolinggo	Wisata Alam Pegunungan Dan Air Terjun	Air Terjun Madakaripura merupakan destinasi wisata alam di Kabupaten Probolinggo yang berada dalam kawasan Taman Nasional Bromo Tengger Semeru. Air terjun ini memiliki ketinggian dan aliran air yang mengelilingi dinding tebing lembah sempit. Lanskap air terjun dan

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				lingkungan pegunungan menjadi karakter utama kawasan wisata ini.
22	Kebun Binatang SBY	Surabaya	Wisata Buatan Dan Rekreasi	Kebun Binatang Surabaya merupakan destinasi wisata buatan di Kota Surabaya yang berfungsi sebagai kawasan konservasi dan edukasi satwa. Kawasan ini menampilkan berbagai jenis fauna dalam lingkungan yang dikelola untuk keperluan pendidikan dan rekreasi. Karakter destinasi ini menggabungkan aspek wisata, edukasi, dan konservasi satwa.
23	Coban Rondo	Malang	Wisata Alam Pegunungan Dan Air Terjun	Coban Rondo merupakan destinasi wisata air terjun di Kabupaten Malang yang berada di kawasan pegunungan dengan lingkungan alam yang sejuk. Air terjun ini dikenal memiliki akses yang relatif mudah serta dikelilingi area hijau yang masih terjaga. Karakter kawasan ini menjadikannya sebagai salah satu destinasi wisata alam pegunungan di Malang.
24	Djawatan	Banyuwangi	Wisata Hutan Dan Konservasi Alam	Djawatan merupakan kawasan wisata hutan di Kabupaten Banyuwangi yang didominasi oleh pepohonan trembesi berusia ratusan tahun. Kawasan ini memiliki lanskap hutan dengan suasana sejuk dan rindang. Karakter vegetasi hutan yang unik menjadikan Djawatan sebagai destinasi wisata alam berbasis hutan.
25	Teluk Ijo	Banyuwangi	Wisata Bahari	Teluk Ijo merupakan destinasi wisata bahari di Kabupaten Banyuwangi yang berada dalam kawasan Taman Nasional Meru Betiri. Pantai ini memiliki pasir putih dan perairan laut berwarna hijau jernih yang dikelilingi tebing serta vegetasi alami. Kondisi alam yang masih terjaga menjadi ciri utama kawasan teluk ini.
26	Kapas Biru	Lumajang	Wisata Alam Pegunungan Dan Air Terjun	Air Terjun Kapas Biru merupakan destinasi wisata alam di Kabupaten Lumajang yang memiliki aliran air bertingkat dengan latar tebing curam. Air terjun ini berada di kawasan lembah sungai yang berhulu di lereng Gunung Semeru dan dikelilingi lanskap alam pegunungan. Karakter air terjun bertingkat dan lingkungan alamnya menjadi ciri utama kawasan wisata ini.
27	Pantai Kasap	Pacitan	Wisata Bahari	Pantai Kasap merupakan destinasi wisata bahari di Kabupaten Pacitan yang memiliki gugusan pulau-pulau

No	Nama Destinasi	Lokasi	Kategori	Deskripsi
				karang kecil di sekitar garis pantai. Pantai ini memiliki pasir putih dan perairan laut berwarna biru dengan bentang alam pesisir yang terbuka. Karakter lanskap pantai berbukit menjadi ciri khas kawasan wisata ini.

Lampiran II. Data Responden PSSUQ

	Nama (Inisial)	Usia		Nama (Inisial)	Usia
R1	D*****	18 - 22 tahun	R30	Y***** A* A***	18 - 22 tahun
R2	S**	18 - 22 tahun	R31	F*	18 - 22 tahun
R3	I** C***** B**	18 - 22 tahun	R32	A***	23 - 27 tahun
R4	M	18 - 22 tahun	R33	R	18 - 22 tahun
R5	A**	18 - 22 tahun	R34	F*	18 - 22 tahun
R6	P*****	18 - 22 tahun	R35	D***	18 - 22 tahun
R7	F*****	18 - 22 tahun	R36	C*	18 - 22 tahun
R8	E***	18 - 22 tahun	R37	D	23 - 27 tahun
R9	A	18 - 22 tahun	R38	A	18 - 22 tahun
R10	R**	18 - 22 tahun	R39	R***	18 - 22 tahun
R11	S***	18 - 22 tahun	R40	A	18 - 22 tahun
R12	N	18 - 22 tahun	R41	F***** F***** N**	23 - 27 tahun
R13	F*****	23 - 27 tahun	R42	L***	18 - 22 tahun
R14	I***	18 - 22 tahun	R43	S*	18 - 22 tahun
R15	L***	18 - 22 tahun	R44	A***	23 - 27 tahun
R16	D	18 - 22 tahun	R45	E	18 - 22 tahun
R17	R*****	18 - 22 tahun	R46	B	18 - 22 tahun
R18	P	18 - 22 tahun	R47	N***	18 - 22 tahun
R19	B**	18 - 22 tahun	R48	L	18 - 22 tahun
R20	F*****	18 - 22 tahun	R49	B	23 - 27 tahun
R21	G*****	23 - 27 tahun	R50	A	18 - 22 tahun
R22	K*****	18 - 22 tahun	R51	Z*****	18 - 22 tahun
R23	W*****	18 - 22 tahun	R52	D***	18 - 22 tahun
R24	N*****	18 - 22 tahun			
R25	O*	18 - 22 tahun			
R26	A*****	18 - 22 tahun			
R27	R	23 - 27 tahun			
R28	I	23 - 27 tahun			
R29	N*	18 - 22 tahun			

Lampiran III. Skenario Uji Coba

No	Destinasi/ <i>Input</i>	Nilai Similarity				
		Rekomendasi 1	Rekomendasi 2	Rekomendasi 3	Rekomendasi 4	Rekomendasi 5
SKENARIO 1						
1	Gunung Bromo	Madakaripura 0.405	Ranu Kumbolo 0.374	Kawah Ijen 0.316	Coban Talun 0.276	Kapas Biru 0.254
2	Tumpak Sewu	Kapas Biru 0.636	Madakaripura 0.559	Coban Rondo 0.519	Coban Talun 0.507	Ranu Kumbolo 0.335
3	Kawah Ijen	Gunung Bromo 0.316	Kapas Biru 0.236	Coban Talun 0.219	Tumpak Sewu 0.219	Madakaripura 0.219
4	Madakaripura	Kapas Biru 0.608	Tumpak Sewu 0.559	Coban Talun 0.525	Coban Rondo 0.469	Gunung Bromo 0.405
5	Ranu Kumbolo	Madakaripura 0.385	Gunung Bromo 0.374	Kapas Biru 0.368	Coban Talun 0.368	Tumpak Sewu 0.335
SKENARIO 2						
1	Pantai Papuma	Balekambang 0.406	Pasir Putih 0.392	Pantai Kasap 0.385	Pantai Klayar 0.369	Teluk Asmara 0.361
2	Balekambang	Pulau Merah 0.545	Pantai Kasap 0.530	Pantai Klayar 0.516	Pantai Tiga Warna 0.504	Pasir Putih 0.495
3	Teluk Ijo	Teluk Asmara 0.424	Pulau Merah 0.354	Pantai Tiga Warna 0.330	Pantai Kasap 0.319	Pasir Putih 0.242
4	Pasir Putih	Teluk Asmara 0.496	Balekambang 0.495	Pantai Kasap 0.494	Pulau Merah 0.399	Pantai Papuma 0.392
5	Pantai Tiga Warna	Balekambang 0.504	Pantai Kasap 0.410	Pulau Merah 0.379	Teluk Asmara 0.356	Pasir Putih 0.338

No	Destinasi/ <i>Input</i>	Nilai Similarity				
		Rekomendasi 1	Rekomendasi 2	Rekomendasi 3	Rekomendasi 4	Rekomendasi 5
SKENARIO 3						
1	Pantai bersih	Pantai Tiga Warna 0.456	Balekambang 0.453	Pantai Klayar 0.391	Teluk Asmara 0.320	Pantai Kasap 0.303
2	Air terjun	Kapas Biru 0.315	Coban Talun 0.280	Madakaripura 0.260	Coban Rondo 0.227	Tumpak Sewu 0.190
3	Pemandangan gunung	Gunung Bromo 0.126	Kawah Ijen 0.094	Ranu Kumbolo 0.065	Teluk Ijo 0.056	Museum Angkut 0.032
4	Wisata edukasi	Jatim Park 2 0.127	Jatim Park 1 0.086	Museum Angkut 0.041	Gunung Bromo 0.038	Kebun Binatang SBY 0.034
5	Wahana bermain	Coban Rondo 0.060	Djawatan 0.050	Jatim Park 2 0.050	BNS 0.048	Pulau Merah 0.044
SKENARIO 4						
1	Pantai yang tenang dan snorkling	Pantai Tiga Warna 0.337	Pantai Klayar 0.264	Balekambang 0.247	Pantai Kasap 0.225	Teluk Asmara 0.196
2	Jalur ekstrim dan pemandangan indah	Coban Talun 0.059	Ranu Kumbolo 0.057	Madakaripura 0.052	Pantai Klayar 0.031	Pantai Kasap 0.030
3	Pemandangan gunung indah dengan kabut	Gunung Bromo 0.137	Kawah Ijen 0.083	Ranu Kumbolo 0.079	Teluk Ijo 0.052	Madakaripura 0.041
4	Wisata edukasi cocok anak keluarga	Jatim Park 2 0.151	Jatim Park 1 0.106	Pasir Putih 0.084	Gunung Bromo 0.063	Taman Safari Prigen 0.063
5	Tempat wisata banyak wahana seru	Jatim Park 1 0.355	BNS 0.262	Selecta 0.087	Jatim Park 2 0.086	Coban Rondo 0.060

Lampiran IV. Rekap Data PSSUQ

Responden																	Jumlah	Nilai Overall $= \frac{jumlah}{16}$
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16		
R1	1	1	1	2	1	1	2	1	1	1	3	1	3	2	1	1	23	1.4375
R2	2	1	1	2	3	2	1	3	2	2	3	1	1	2	2	2	30	1.875
R3	2	2	2	3	3	2	2	2	2	2	3	2	2	2	2	3	36	2.25
R4	1	3	3	4	1	2	1	2	1	2	2	2	1	1	2	1	29	1.8125
R5	1	1	1	3	1	1	1	1	1	1	3	1	1	1	1	1	20	1.25
R6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	1
R7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	1
R8	1	1	3	3	3	1	2	2	3	4	5	3	1	1	1	3	37	2.3125
R9	4	3	4	4	4	3	1	2	1	3	5	4	4	1	3	3	49	3.0625
R10	1	1	1	2	2	1	2	2	2	1	2	1	1	1	1	2	23	1.4375
R11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	1
R12	3	3	3	2	2	3	1	1	1	1	1	1	1	1	1	1	26	1.625
R13	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	32	2
R14	2	1	1	1	2	2	2	2	1	1	1	1	1	1	2	1	22	1.375
R15	3	4	5	3	3	2	2	4	3	3	3	2	2	2	3	2	46	2.875
R16	2	1	2	2	3	3	2	1	2	1	2	3	3	2	2	3	34	2.125
R17	3	2	1	4	3	3	1	1	1	3	3	2	2	1	1	1	32	2
R18	3	1	1	2	1	4	2	2	2	2	1	3	1	1	2	2	30	1.875
R19	2	1	3	2	2	1	1	2	3	1	1	1	2	2	4	4	32	2
R20	2	2	1	2	1	3	4	3	2	2	1	2	2	2	4	2	35	2.1875
R21	2	2	2	2	2	1	2	2	1	1	2	2	1	1	2	2	27	1.6875
R22	2	2	3	3	3	3	4	3	2	3	3	3	3	3	3	3	46	2.875
R23	3	3	2	2	2	2	2	2	2	3	2	3	2	2	2	2	36	2.25
R24	2	2	1	3	2	3	3	2	1	1	3	2	1	2	3	1	32	2
R25	3	2	2	3	3	4	3	1	2	3	2	2	3	3	3	3	42	2.625
R26	2	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	18	1.125
R27	2	1	2	3	1	2	5	4	3	2	2	3	3	2	1	2	38	2.375

Responden																	Jumlah	Nilai Overall = $\frac{jumlah}{16}$
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16		
R28	3	3	4	2	3	4	3	3	3	3	4	3	3	4	3	4	52	3.25
R29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	1
R30	1	1	1	2	1	1	4	3	3	3	2	5	2	2	3	2	36	2.25
R31	1	2	3	1	3	2	3	3	3	2	2	3	1	1	3	2	35	2.1875
R32	4	4	4	4	3	3	4	3	3	2	3	2	3	2	3	2	49	3.0625
R33	4	2	3	4	2	4	4	4	6	4	5	5	4	4	4	4	63	3.9375
R34	2	2	1	2	1	1	1	1	1	3	2	3	2	1	1	2	26	1.625
R35	2	1	2	1	2	1	2	3	2	2	2	2	1	1	2	2	28	1.75
R36	2	1	1	1	2	1	2	3	1	2	1	2	1	1	1	1	23	1.4375
R37	2	2	2	2	2	2	2	2	2	2	2	2	6	6	3	5	44	2.75
R38	2	1	2	1	2	1	2	2	1	2	1	1	1	1	1	1	22	1.375
R39	4	2	3	3	3	3	3	3	3	3	2	3	2	2	2	2	43	2.6875
R40	3	2	2	3	1	3	4	3	4	4	3	1	4	4	4	3	48	3
R41	4	3	2	3	2	1	3	3	3	3	2	3	3	3	3	3	44	2.75
R42	2	1	2	2	3	2	2	2	3	2	2	3	1	1	2	2	32	2
R43	2	2	1	1	2	1	3	2	1	1	1	2	2	2	2	2	27	1.6875
R44	2	1	1	2	2	3	4	3	2	2	3	1	3	2	3	1	35	2.1875
R45	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4	3	35	2.1875
R46	2	1	2	2	1	2	4	4	2	2	2	1	3	3	2	2	35	2.1875
R47	4	3	3	3	4	3	3	4	3	3	3	3	4	4	4	3	54	3.375
R48	2	1	2	2	1	2	3	3	2	1	2	2	2	2	2	2	31	1.9375
R49	2	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	19	1.1875
R50	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	49	3.0625
R51	2	1	2	2	2	3	4	3	2	2	1	3	4	4	3	2	40	2.5
R52	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	1
Rata-rata																		2.073

