

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan platform ulasan berbasis lokasi, khususnya Google Maps, telah mengubah cara wisatawan menilai kualitas destinasi dan mengambil keputusan berkunjung karena ulasan memuat pengalaman aktual, persepsi layanan, dan evaluasi fasilitas dalam bentuk teks serta rating numerik. Kombinasi teks dan rating menjadikan ulasan digital bukan hanya media ekspresi pengalaman pribadi, tetapi juga sumber informasi publik yang membentuk ekspektasi calon pengunjung (Maududie et al., 2024). Dalam konteks wisata keluarga dan edukasi, Kebun Binatang Surabaya (KBS) menjadi objek yang relevan karena melibatkan beragam segmen pengunjung dan menghasilkan umpan balik yang kaya tentang pengalaman kunjungan (Merawati Putu et al., 2021). Ulasan pengunjung umumnya mencerminkan pengalaman pra-kunjungan, saat kunjungan, hingga pasca-kunjungan, seperti harga tiket, akses, parkir, kebersihan, kenyamanan, atraksi, hingga kepuasan umum yang direpresentasikan melalui rating dan narasi teks (Uyun & Nuryana, 2025). Dari sisi analitik, teks ulasan dapat digunakan untuk menemukan tema pengalaman dominan, sedangkan rating dapat dimanfaatkan sebagai dasar segmentasi kepuasan sehingga keduanya dapat memotret persepsi publik secara lebih komprehensif (Zahrah et al., 2024).

Namun, ulasan digital juga menyimpan masalah bagi pengelolaan destinasi karena jumlah ulasan yang besar, ragam bahasa yang tinggi, dan bentuk teks yang tidak terstruktur membuat isu inti sulit diidentifikasi secara cepat dan objektif (Aryandi et al., 2023). Satu ulasan sering memuat campuran beberapa topik, seperti akses, fasilitas, satwa, wahana berbayar, kepadatan, dan pedagang, sehingga pembacaan manual berisiko melewatkan pola penting yang tersebar pada banyak ulasan. Kondisi ini membuat pengambilan keputusan berbasis pembacaan manual rawan bias, terutama ketika pengelola lebih mudah mengingat ulasan ekstrem daripada pola berulang yang tersebar (Sibarani et al., 2024). Kompleksitas semakin tinggi karena rating dan isi teks tidak selalu konsisten, misalnya rating tinggi yang masih memuat keluhan tertentu atau rating sedang yang tetap memuji aspek spesifik, sehingga skor agregat saja tidak cukup untuk menentukan prioritas

perbaikan (Uyun & Nuryana, 2025). Oleh karena itu, pengelolaan ulasan destinasi memerlukan metode analitik yang mampu memetakan tema secara sistematis, terukur, dan relevan untuk perbaikan layanan (Aryandi et al., 2023).

Berdasarkan kondisi tersebut, penelitian ini memfokuskan pada pemodelan topik utama ulasan Google Maps KBS serta mengaitkannya dengan segmentasi rating pengunjung agar isu dan keunggulan layanan dapat dipetakan lebih tajam menurut tingkat kepuasan (Sibarani et al., 2024). Segmentasi rating penting karena tema pada rating tinggi, seperti “murah”, “rindang”, atau “ramah anak”, dapat menunjukkan faktor pendorong kepuasan yang perlu dipertahankan, sedangkan tema pada rating rendah, seperti “kondisi kandang”, “bau”, “kebersihan”, atau “kepadatan”, dapat menunjukkan sumber keluhan prioritas yang perlu segera ditindaklanjuti (Zahrah et al., 2024). Topik yang sama juga dapat memiliki makna berbeda pada segmen rating yang berbeda. Misalnya, kepadatan dan parkir masih dapat ditoleransi pada rating tinggi, tetapi menjadi sumber ketidaknyamanan utama pada rating rendah (Uyun & Nuryana, 2025). Hal ini menunjukkan bahwa analisis tanpa segmentasi rating berisiko menghasilkan interpretasi yang terlalu umum dan kurang tepat untuk penetapan prioritas layanan. Karena itu, penelitian ini memerlukan pendekatan kuantitatif yang mampu mengekstraksi pola topik dominan sekaligus membandingkan distribusinya antar segmen rating agar hasilnya lebih operasional bagi pengelola KBS (Merawati Putu et al., 2021).

Sejumlah pendekatan telah digunakan untuk mengolah ulasan wisata, seperti analisis sentimen berbasis klasifikasi terawasi, analisis frekuensi kata, dan pemodelan topik untuk mengekstraksi tema laten dari kumpulan dokumen (Zahrah et al., 2024). Pada ranah wisata berbasis Google Maps, text mining telah digunakan untuk merangkum masukan pengunjung menjadi informasi yang lebih mudah ditindaklanjuti (Maududie et al., 2024). Pemodelan topik dengan Latent Dirichlet Allocation (LDA) juga telah banyak diterapkan pada ulasan destinasi karena mampu memetakan kata-kata yang sering muncul bersama menjadi topik yang representatif (Sibarani et al., 2024). Selain di bidang wisata, LDA juga digunakan pada ulasan layanan untuk mengekstraksi faktor ketidakpuasan, sehingga menunjukkan fleksibilitas metode ini pada berbagai domain (Uyun & Nuryana, 2025). Dalam konteks penelitian ini, LDA dengan segmentasi rating dipilih karena

dapat menjembatani kebutuhan identifikasi tema dominan dan kebutuhan perbandingan antar tingkat kepuasan, sehingga hasil analisis lebih tajam untuk rekomendasi perbaikan dan pemeliharaan layanan (Sibarani et al., 2024).

Pendekatan lain seperti klasifikasi sentimen terawasi memang berguna untuk membedakan ulasan positif dan negatif, tetapi umumnya membutuhkan data berlabel yang memadai dan konsisten, sementara proses pelabelan manual sering memakan waktu dan berpotensi subjektif (Zahrah et al., 2024). Analisis frekuensi kata lebih mudah diterapkan, tetapi kurang mampu menangkap konteks dan hubungan antar kata sehingga hasilnya sering terlalu dangkal untuk kebutuhan operasional (Aryandi et al., 2023). Sementara itu, LDA unggul untuk mengeksplorasi struktur tema tanpa label dan merangkum korpus besar menjadi sejumlah topik, meskipun kualitas hasilnya tetap dipengaruhi oleh praproses, jumlah topik, dan kebisingan data (Suhendro & Dora, 2025). Selain itu, beberapa studi pemodelan topik pada ulasan destinasi masih menyajikan topik secara agregat, sehingga informasi tentang perbedaan topik pada rating rendah dan rating tinggi belum dibahas secara mendalam sebagai dasar prioritas layanan (Suhendro & Dora, 2025).

Keterbatasan penelitian sebelumnya menunjukkan bahwa meskipun LDA mampu mengekstraksi tema layanan dan fasilitas dari ulasan wisata, banyak studi masih berhenti pada interpretasi topik global dan belum menekankan segmentasi rating sebagai pembanding sistematis antar kelompok kepuasan (Uyun & Nuryana, 2025). Pada konteks Google Maps, studi text mining telah membantu merangkum masukan pengunjung, tetapi ruang kontribusi masih terbuka pada desain analisis kuantitatif yang menghubungkan distribusi topik dengan rating secara eksplisit (Maududie et al., 2024). Kesenjangan ini menjadi penting pada konteks KBS karena destinasi ini berskala besar, memiliki layanan yang beragam, dan berpotensi menghasilkan isu yang heterogen dengan tingkat kepentingan berbeda menurut pengalaman pengunjung (Merawati Putu et al., 2021). Berdasarkan kesenjangan tersebut, penelitian ini menawarkan pemodelan topik ulasan Google Maps KBS menggunakan Latent Dirichlet Allocation (LDA) dengan segmentasi berdasarkan rating pengunjung (Uyun & Nuryana, 2025). Segmentasi rating rendah, sedang, dan tinggi dipilih karena rating merupakan indikator kuantitatif yang langsung tersedia

pada platform ulasan dan dapat dimanfaatkan tanpa menambah beban pelabelan manual (Suhendro & Dora, 2025). Melalui pendekatan ini, topik dominan pada rating rendah dapat ditafsirkan sebagai sumber keluhan prioritas, sedangkan topik dominan pada rating tinggi dapat ditafsirkan sebagai kekuatan layanan yang perlu dipertahankan. Dengan demikian, penelitian ini diharapkan menghasilkan peta tema yang lebih tersegmentasi, lebih tajam, dan lebih relevan untuk pengambilan keputusan berbasis bukti pada pengelolaan KBS (Zahrah et al., 2024).

1.2 Perumusan Masalah

Berdasarkan latar belakang penelitian ini, rumusan masalah yang diuraikan sebagai berikut :

1.2.1 Permasalahan

Secara ilmiah dari sudut pandang Informatika, permasalahan pada penelitian ini adalah belum optimalnya proses ekstraksi informasi tematik dari data ulasan Google Maps yang bersifat teks tidak terstruktur, bising, dan heterogen, sehingga sulit menghasilkan pemetaan isu layanan yang terukur untuk mendukung pengambilan keputusan. Selain itu, analisis ulasan yang dilakukan secara agregat cenderung menutupi perbedaan tema antara kelompok pengunjung dengan tingkat kepuasan berbeda, padahal rating dapat digunakan sebagai sinyal segmentasi untuk meningkatkan ketajaman analisis topik. Akibatnya, pengelola berisiko tidak memperoleh prioritas perbaikan layanan yang spesifik berdasarkan pola keluhan dan kepuasan pengunjung.

1.2.2 Metode Usulan

Metode yang diusulkan untuk mengatasi permasalahan tersebut adalah pemodelan topik menggunakan Latent Dirichlet Allocation (LDA) pada data ulasan Google Maps Kebun Binatang Surabaya, dengan segmentasi berdasarkan rating pengunjung. Pada penelitian ini, rating digunakan sebagai dasar segmentasi menjadi tiga kelompok, yaitu rating rendah, rating sedang, dan rating tinggi, untuk memisahkan kelompok ulasan sebelum dilakukan pemodelan topik. Pendekatan ini didukung oleh tahapan *preprocessing* teks bahasa Indonesia, pembentukan representasi dokumen, pemodelan LDA per segmen, serta evaluasi kualitas topik menggunakan metrik koherensi topik untuk memperoleh hasil yang lebih interpretatif dan relevan bagi rekomendasi layanan.

1.2.3 Pertanyaan Penelitian

1. Apa saja topik dominan yang dihasilkan oleh metode Latent Dirichlet Allocation (LDA) pada ulasan Google Maps Kebun Binatang Surabaya untuk setiap segmen rating pengunjung?
2. Berapa nilai koherensi topik terbaik dan bagaimana peran segmentasi rating pengunjung terhadap ketajaman hasil pemodelan topik LDA dalam mengidentifikasi isu keluhan dan faktor kepuasan layanan?

1.3 Tujuan & Manfaat

Subbab ini menjelaskan tujuan yang ingin dicapai penelitian dan manfaat yang diharapkan ketika tujuan tersebut terealisasi. Tujuan bersifat spesifik dan sejalan dengan judul penelitian, sedangkan manfaat menyoroti kontribusi teoritis maupun praktis dari hasil penelitian.

1.3.1 Tujuan Penelitian

1. Mengidentifikasi topik-topik dominan yang dihasilkan oleh metode Latent Dirichlet Allocation (LDA) pada ulasan Google Maps Kebun Binatang Surabaya untuk setiap segmen rating pengunjung.
2. Menentukan nilai koherensi topik terbaik serta menganalisis peran segmentasi rating pengunjung terhadap ketajaman hasil pemodelan topik LDA dalam mengidentifikasi isu keluhan dan faktor kepuasan layanan.

1.3.2 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memperkaya kajian *text mining* dan *topic modeling* pada data ulasan digital berbasis lokasi, khususnya Google Maps, dalam konteks destinasi wisata dan rekreasi keluarga. Kontribusi teoritis utama penelitian ini terletak pada penggunaan segmentasi rating sebagai pendekatan analitik untuk membaca perbedaan struktur topik berdasarkan tingkat kepuasan pengunjung, sehingga analisis tidak berhenti pada topik agregat.

2. Manfaat Praktis

- a. Bagi pengelola Kebun Binatang Surabaya

Hasil penelitian dapat digunakan untuk mengidentifikasi tema keluhan dominan dan tema kepuasan dominan secara lebih terstruktur, sehingga membantu penetapan prioritas perbaikan layanan.

b. Bagi manajemen operasional destinasi

Pemetaan topik per segmen rating dapat menjadi dasar evaluasi aspek layanan yang perlu dipertahankan dan aspek yang perlu ditindak lanjuti segera.

c. Bagi Ilmu Pengetahuan

Rancangan penelitian ini dapat menjadi referensi metodologis untuk pengembangan studi lanjutan, misalnya analisis temporal topik, perbandingan metode topic modeling, atau kombinasi topic modeling dengan analisis sentimen.

1.4 Batasan - Batasan Penelitian

Agar penelitian fokus dan terarah, beberapa batasan yang ditetapkan adalah sebagai berikut:

1. Objek penelitian adalah ulasan pengunjung Kebun Binatang Surabaya (KBS) yang tersedia pada platform Google Maps, dengan fokus pada konten ulasan berbentuk teks dan rating bintang yang menyertainya.
2. Sumber data yang digunakan hanya berasal dari Google Maps, sehingga penelitian ini tidak mencakup ulasan dari platform lain.
3. Variabel utama yang dianalisis meliputi Teks ulasan sebagai data utama untuk pemodelan topik dan Rating pengunjung sebagai dasar segmentasi kelompok ulasan.
4. Segmentasi rating dilakukan berdasarkan tiga kelompok:
 - a. Rating rendah (1–2) sebagai representasi kecenderungan keluhan.
 - b. Rating sedang (3) sebagai representasi penilaian sedang, yaitu ulasan yang tidak sepenuhnya puas maupun tidak puas.
 - c. Rating tinggi (4–5) sebagai representasi kecenderungan kepuasan.

Rating digunakan sebagai indikator segmentasi kecenderungan, bukan sebagai kebenaran absolut polaritas isi teks.

5. Analisis dilakukan dengan membandingkan topik dominan pada ketiga segmen rating untuk melihat perbedaan tema keluhan, tema evaluasi, dan tema kepuasan secara lebih terstruktur.
6. Analisis penelitian berfokus pada pemodelan topik (topic modeling) menggunakan LDA, sehingga penelitian ini tidak melakukan klasifikasi sentimen terawasi (misalnya SVM, Naive Bayes, atau model deep learning) sebagai metode utama.
7. Bahasa ulasan yang dianalisis dibatasi pada ulasan berbahasa Indonesia, agar hasil pemodelan topik lebih konsisten dan interpretasi topik lebih akurat dalam konteks penelitian.
8. Periode data ulasan dibatasi pada rentang waktu yang ditetapkan saat pengumpulan data penelitian yaitu 1 tahun terakhir, sehingga hasil penelitian merepresentasikan kondisi ulasan pada periode tersebut.

1.5 Sistematika Penulisan

Penelitian disusun dalam beberapa bab agar alur pembahasan jelas dan sistematis, dengan uraian sebagai berikut:

1. Bab I Pendahuluan menjelaskan latar belakang penelitian, perumusan masalah, tujuan dan manfaat penelitian, batasan-batasan serta sistematika penulisan.
2. Bab II Tinjauan Pustaka Bab ini berisi landasan teoritis dan empiris yang mendukung penelitian. Pembahasan difokuskan pada teori, konsep, dan hasil penelitian terdahulu yang relevan dengan pemodelan topik ulasan digital, LDA, serta analisis berbasis rating.
3. Bab III Metodologi Penelitian menjelaskan rancangan teknis penelitian secara rinci agar proses penelitian dapat dipahami, diulang, dan dievaluasi secara ilmiah, bab ini menitikberatkan pada prosedur pengolahan data ulasan, segmentasi rating, dan pemodelan topik menggunakan LDA.
4. BAB IV Hasil dan Pembahasan Bab ini menyajikan hasil penelitian dan pembahasan dari setiap tahap pengolahan data. Pembahasan meliputi deskripsi data, pengelompokan rating, preprocessing teks, pembentukan bigram dan trigram, Bag of Words, hasil pemodelan topik dengan LDA,

visualisasi wordcloud dan pyLDAvis, uji coba skenario, interpretasi topik, tampilan sistem, serta perbandingan dengan penelitian sebelumnya.

5. BAB V Penutup Bab ini berisi kesimpulan dan saran berdasarkan hasil penelitian. Kesimpulan disusun untuk menjawab pertanyaan penelitian, sedangkan saran diberikan sebagai masukan untuk penelitian selanjutnya dan pengembangan analisis ulasan pengunjung pada masa mendatang.