

**PENAMBANGAN DATA ONLINE UNTUK PENGEMBANGAN
SEKTOR PARIWISATA DENGAN MENINGTEGRASIKAN METODE
LATENT DIRICHLET ALLOCATION (LDA), ANALISIS SENTIMEN,
DAN IMPORTANCE-PERFORMANCE ANALYSIS (IPA)
STUDI KASUS: SEKTOR PARIWISATA YOGYAKARTA GOA
PINDUL**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk memperoleh Gelar Sarjana Strata-1
Program Studi teknik industri – Fakultas Teknologi Industri
Universitas Islam Indonesia**



Nama : Reza Ibrahim
No. Mahasiswa : 22522241

**PROGRAM STUDI TEKNIK INDUSTRI PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA**

2026

PERNYATAAN KEASLIAN

Saya mengakui bahwa tugas akhir ini adalah hasil karya saya sendiri kecuali kutipan dan ringkasan yang seluruhnya sudah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah maka saya bersedia ijazah yang telah saya terima ditarik kembali oleh Universitas Islam Indonesia.

Yogyakarta, 26 – 05 – 2026



(Reza Ibrahim)
22522241

SURAT BUKTI PENELITIAN**FAKULTAS
TEKNOLOGI INDUSTRI**

Gedung KH. Mas Mansur
Kampus Terpadu Universitas Islam Indonesia
Jl. Kalurungkm 14,5 Yogyakarta 55584
T. (0274) 898444 ext. 4100, 4101
F. (0274) 895007
E. ti@uii.ac.id
W. tiuii.ac.id

SURAT KETERANGAN PENELITIAN

Nomor: 018/Ka.Lab.Datamin/70/Lab.Datamin/V/2026

Assalamu'alaikum Warahmatullahi Wabarakatuh

Kami yang bertanda tangan dibawah ini, menerangkan bahwa mahasiswa dengan keterangan sebagai berikut :

Nama : Reza Ibrahim
No. Mhs : 22522241
Dosen Pembimbing : Dr. Dwi Adi Purnama, S.T.

Telah selesai melaksanakan penelitian yang berjudul "Penambangan Data Online untuk Pengembangan Sektor Pariwisata dengan Mengintegrasikan Metode Latent Dirichlet Allocation (LDA), Analisis Sentimen, dan Importance-Performance Analysis (IPA)" di Laboratorium Data Mining, Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 01 Maret sampai dengan tanggal 08 April 2026.

Demikian surat keterangan kami keluarkan, agar dapat dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 29 Mei 2026

Kepala Laboratorium
Data Mining

Dr. Dwi Adi Purnama, S.T.

LEMBAR PENGESAHAN PEMBIMBING

**PENAMBANGAN DATA ONLINE UNTUK PENGEMBANGAN SEKTOR
PARIWISATA DENGAN MENGINTEGRASIKAN METODE LATENT DIRICHLET
ALLOCATION (LDA), ANALISIS SENTIMEN, DAN IMPORTANCE-
PERFORMANCE ANALYSIS (IPA)
STUDI KASUS: SEKTOR PARIWISATA YOGYAKARTA GOA PINDUL**



Yogyakarta, 26 Mei 2026

Dosen Pembimbing

A handwritten signature in black ink, consisting of several loops and strokes, representing the signature of Dr. Dwi Adi Purnama, S.T.

Dr. Dwi Adi Purnama, S.T.

LEMBAR PENGESAHAN DOSEN PENGUJI

PENAMBANGAN DATA ONLINE UNTUK PENGEMBANGAN SEKTOR
PARIWISATA DENGAN MENINGTEGRASIKAN METODE LATENT DIRICHLET
ALLOCATION (LDA), ANALISIS SENTIMEN, DAN IMPORTANCE-
PERFORMANCE ANALYSIS (IPA)
STUDI KASUS: SEKTOR PARIWISATA YOGYAKARTA GOA PINDUL

TUGAS AKHIR

Disusun Oleh :

Nama : Reza Ibrahim
 No. Mahasiswa : 22522241

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk
 memperoleh gelar Sarjana Strata-1 Teknik Industri Fakultas Teknologi Industri
 Universitas Islam Indonesia

Yogyakarta, 15 Juni 2026

Tim Penguji

Dr. Dwi Adi Purnama, S.T.
 Ketua

Dr. Harwati, S.T., M.T.
 Anggota I

Didin Dwi Novianto, S.T., M.LSCM.
 Anggota II

Mengetahui,

Ketua Program Studi Teknik Industri Program Sarjana
 Fakultas Teknologi Industri
 Universitas Islam Indonesia

Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM

NIK. 015220101

HALAMAN PERSEMBAHAN

Tugas Akhir ini dipersembahkan kepada orang tua dan keluarga tercinta atas doa, dukungan, serta pengorbanan yang tiada henti. Penulis juga menyampaikan penghargaan kepada dosen pembimbing dan seluruh sivitas akademika Universitas Islam Indonesia atas bimbingan serta ilmu yang diberikan selama proses penyusunan.

MOTTO

“Maka sesungguhnya beserta kesulitan ada kemudahan.
Sesungguhnya beserta kesulitan ada kemudahan”
(QS. Al-Insyirah: 5-6)

“Sesungguhnya Allah tidak mengubah keadaan suatu kaum hingga
mereka mengubah keadaan yang ada pada diri mereka”
(QS. Ar-Ra’d: 11)

*“If you are not a good shot today, don’t worry.
There are other ways to be useful”*

KATA PENGANTAR

Puji syukur Penulis panjatkan ke hadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga Penulis dapat menyelesaikan Tugas Akhir ini dengan baik. Penulis tentu menyadari, bahwa Tugas Akhir ini tidak akan diselesaikan tanpa adanya dukungan dan bantuan dari berbagai pihak. Oleh karena itu, Penulis berterima kasih kepada semua pihak yang telah memberikan bantuan dan dukungan dalam penyusunan laporan ini. Pada kesempatan ini, Penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo., M.T., IPU, ASEAN.Eng selaku Dekan Fakultas Teknologi Industri Universitas Islam Indonesia
2. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM selaku Ketua Program Studi Teknik Industri
3. Bapak Dr. Dwi Adi Purnama, S.T. selaku dosen pembimbing Tugas Akhir yang telah memberikan bimbingan, arahan, serta berbagai masukan sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan baik.
4. Kepada kedua orang tua dan saudara kandung Penulis yang selalu mendoakan dan memberikan semangat dalam bentuk apapun yang tiada hentinya.
5. Kepada teman-teman terbaik yang tergabung dalam *group* “LS” yang telah menjadi bagian tak terpisahkan dalam perjalanan panjang ini. Sejak masa putih abu-abu hingga menapaki akhir bangku perkuliahan, kalian hadir bukan sekadar sebagai teman, melainkan sebagai rumah tempat berbagi cerita. Canda dan tawa yang tak pernah habis, serta kebersamaan dalam suka maupun duka, telah menjadi warna yang menguatkan setiap langkah Penulis. Dalam setiap proses yang dilalui, kehadiran kalian menjadi pengingat bahwa perjalanan ini tidak pernah dijalani sendirian.
6. Kepada teman-teman saya yaitu Abiyyu, Agim, Alid, Allan, Arul, Cherie, Duta, Fadil, Fadlan, Gean, Iyan, Nandes, Nasywa, NW, Raka, Regung, dan Ryan yang telah menemani penulis berjuang dalam menempuh perjalanan selama di masa perkuliahan yang tidak pernah terlupakan. Tanpa dukungan, semangat, canda-tawa, dan cerita yang diberikan, penulis belum tentu dapat menyelesaikan masa studi ini dengan lancar dan baik. Terima kasih juga atas slogan yang pernah dibuat bersama dengan teman-teman Teknik Industri FTI UII Angkatan 2022, yaitu “Bersama-sama Mengukir Cerita”.

ABSTRAK

Perkembangan sektor pariwisata di Indonesia yang terus meningkat menuntut pengelolaan destinasi wisata yang lebih adaptif dan berbasis data. Wisata Goa Pindul sebagai salah satu destinasi unggulan di Daerah Istimewa Yogyakarta menghadapi tantangan dalam menjaga kualitas layanan seiring meningkatnya jumlah pengunjung. Penelitian ini bertujuan untuk mengidentifikasi topik utama dalam ulasan wisatawan, menganalisis kecenderungan sentimen, serta menentukan prioritas perbaikan layanan sebagai dasar rekomendasi strategi pengembangan wisata. Data penelitian diperoleh melalui teknik web scraping dari *platform Google Maps Review* dan *Instagram* pada periode Januari 2020 hingga Februari 2026. Metode yang digunakan meliputi *Latent Dirichlet Allocation* (LDA) untuk pemodelan topik, analisis sentimen menggunakan VADER untuk mengidentifikasi polaritas opini, serta *Importance Performance Analysis* (IPA) untuk menentukan prioritas perbaikan atribut layanan.

Hasil penelitian menunjukkan bahwa ulasan wisatawan didominasi oleh beberapa topik utama seperti fasilitas, pelayanan pemandu, harga, aksesibilitas, serta pengalaman wisata. Analisis sentimen menunjukkan kecenderungan opini yang didominasi sentimen positif, namun masih terdapat sentimen negatif pada aspek tertentu seperti kebersihan dan fasilitas pendukung. Berdasarkan hasil IPA, atribut yang masuk dalam prioritas utama perbaikan meliputi aspek pelayanan, pengelolaan fasilitas, dan kepuasan pengunjung. Penelitian ini menghasilkan rekomendasi strategis yang dapat digunakan oleh pengelola untuk meningkatkan kualitas layanan dan kepuasan wisatawan secara berkelanjutan.

Kata Kunci: LDA, Analisis Sentimen, VADER, *Importance Performance Analysis*, Pariwisata

DAFTAR ISI

PERNYATAAN KEASLIAN	ii
SURAT BUKTI PENELITIAN	iii
LEMBAR PENGESAHAN PEMBIMBING	iv
LEMBAR PENGESAHAN DOSEN PENGUJI	v
HALAMAN PERSEMBAHAN	vi
MOTTO	vii
KATA PENGANTAR	viii
ABSTRAK	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Kajian Literatur	7
2.1.1. Penelitian Web Scrapping pada Pariwisata	7
2.1.2. Penelitian Analisis Sentimen dalam Pariwisata	8
2.1.3. Penelitian Topic Modelling Menggunakan Latent Dirichlet Allocation (LDA)	8
2.1.4. Penelitian Importance Analysis (IPA) dalam Evaluasi Destinasi.....	9
2.1.5. Research Gap dan Posisi Penelitian	9
2.2 Landasan Teori	14
2.2.1. Pariwisata	14
2.2.2. Data Scrapping	14
2.2.3. Python.....	14
2.2.4. Google Maps Review	15
2.2.5. Instagram	15
2.2.6. Text Mining	15
2.2.7. Latent Dirichlet Allocation (LDA).....	16
2.2.8. Analisis Sentimen.....	19
2.2.9. Valence Aware Dictionary and sEntiment Reasoner (VADER)	20
2.2.10. Importance-Performance Analysis (IPA)	22
BAB III METODOLOGI PENELITIAN	25
3.1 Subjek Penelitian.....	25
3.2 Objek Penelitian	25
3.3 Sumber dan Jenis Data	25
3.4 Alur Penelitian.....	25
BAB IV PENGOLAHAN DATA	31
4.1 Pengumpulan Data	31
4.1.1. Data Scrapping Google Maps	31
4.1.2. Data Scrapping Instagram	31
4.2 Preprocessing Data	31

4.2.1.	Translasi	31
4.2.2.	Filtering	32
4.2.3.	Case Folding.....	34
4.2.4.	Stopwords Removal	35
4.2.5.	Tokenisasi.....	36
4.2.6.	Stemming	37
4.3	<i>Latent Dirichlet Allocation (LDA)</i>	38
4.3.1.	Pemodelan Topik Data Google Maps.....	38
4.3.2.	Pemodelan Topik Data Instagram	40
4.4	<i>Valence Aware Dictionary and sEntiment Reseanor (VADER)</i>	42
4.4.1.	Analisis Sentimen Data Google Maps	42
4.4.2.	Analisis Sentimen Data Instagram	43
4.5	<i>Importance-Performance Analysis (IPA) Matrix</i>	44
4.5.1.	Penentuan Nilai Importance	44
4.5.2.	Penentuan Nilai Performance	46
4.5.3.	Visualisasi Importance-Performance Analysis (IPA) Matrix	47
BAB V	HASIL DAN PEMBAHASAN	50
5.1.	Hasil <i>Latent Dirichlet Allocation (LDA)</i>	50
5.2.	Hasil Analisis Sentimen	55
5.3.	Hasil <i>Importance-Performance Analysis (IPA) Matrix</i>	57
5.3.1.	Google Maps IPA Matrix	58
5.3.2.	Instagram IPA Matrix	60
5.4.	Rekomendasi Strategi Pengembangan Wisata Goa Pindul	61
5.4.1.	Perancangan Sistem Audit Kebersihan dan Pengendalian Lingkungan Wisata 62	
5.4.2.	Perancangan Customer Journey Wisata Keluarga.....	63
5.4.3.	Perancangan Dashboard Voice of Customer (VOC)	66
5.4.4.	Perancangan SOP dan Standarisasi Pelayanan Admin.....	67
BAB VI	PENUTUP	71
6.1.	Kesimpulan.....	71
6.2.	Saran.....	72
DAFTAR	PUSTAKA	74
LAMPIRAN		A-1

DAFTAR TABEL

Tabel 2. 1 Posisi penelitian	11
Tabel 2. 2 Perbandingan metode pemodelan topik.....	19
Tabel 2. 3 Perbandingan metode analisis sentimen	21
Tabel 3. 1 Klasifikasi sentimen.....	29
Tabel 4. 1 Hasil <i>text</i> translasi.....	32
Tabel 4. 2 <i>Function filtering Excel</i>	33
Tabel 4. 3 Hasil <i>filtering</i>	33
Tabel 4. 4 Hasil <i>case folding</i>	34
Tabel 4. 5 <i>Function Excel</i> untuk <i>stopwords removal</i>	35
Tabel 4. 6 Hasil <i>stopword removal</i>	36
Tabel 4. 7 Hasil tokenisasi.....	37
Tabel 4. 8 Hasil <i>stemming</i>	37
Tabel 4. 9 <i>Label</i> topik ulasan <i>Google Maps</i>	40
Tabel 4. 10 <i>Label</i> topik ulasan <i>Instagram</i>	42
Tabel 4. 11 <i>Sentiment score</i> ulasan <i>Google Maps</i>	43
Tabel 4. 12 <i>Sentimen score</i> ulasan <i>Instagram</i>	43
Tabel 4. 13 Total bobot kepentingan topik ulasan	44
Tabel 4. 14 <i>Importance score</i> topik ulasan <i>Google Maps</i>	45
Tabel 4. 15 <i>Performance score</i> IPA <i>matrix</i>	46
Tabel 4. 16 <i>Performance</i> dan <i>importance score</i>	47
Tabel 5. 1 Interpretasi topik pada ulasan <i>Google Maps</i>	50
Tabel 5. 2 Interpretasi topik pada ulasan <i>Instagram</i>	52
Tabel 5. 3 Jumlah <i>label</i> analisis sentimen ulasan <i>Google Maps</i>	56
Tabel 5. 4 Jumlah <i>label</i> sentimen ulasan <i>Instagram</i>	56
Tabel 5. 5 Sistem Audit Kebersihan	62
Tabel 5. 6 Perancangan <i>Customer Journey</i> Wisata Keluarga	64
Tabel 5. 7 SOP dan Standarisasi Pelayanan Admin.....	68

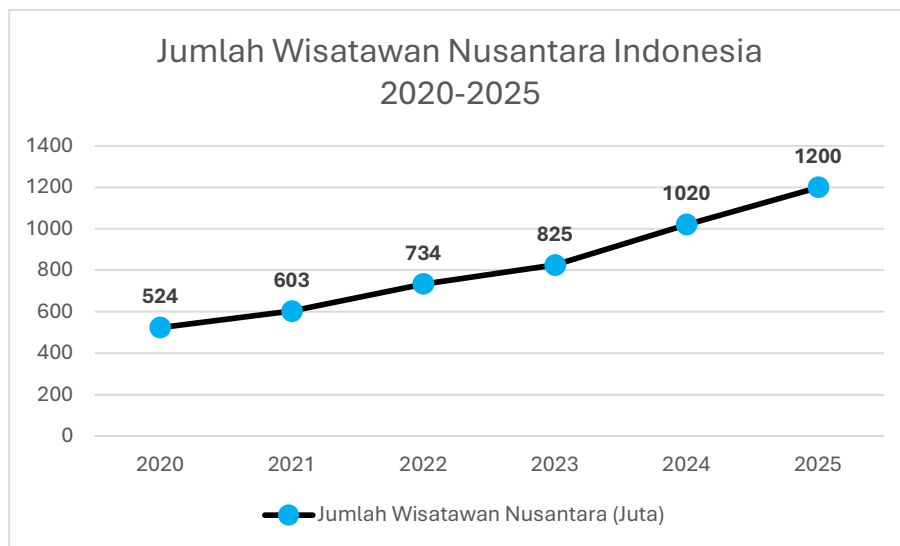
DAFTAR GAMBAR

Gambar 1. 1 Perkembangan pariwisata di Indonesia	1
Gambar 1. 2 Perkembangan pariwisata di D.I. Yogyakarta.....	2
Gambar 1. 3 Perkembangan jumlah pengunjung Goa Pindul tahunan.....	3
Gambar 2. 1 Alur kerja metode LDA	17
Gambar 2. 2 Pengelompokan metode analisis sentimen	20
Gambar 2. 3 IPA <i>matrix</i>	22
Gambar 3. 1 Alur penelitian	26
Gambar 4. 1 Grafik koherensi LDA ulasan <i>Google Maps</i>	39
Gambar 4. 2 Grafik koherensi LDA ulasan <i>Instagram</i>	41
Gambar 4. 3 IPA <i>matrix</i> ulasan Goa Pindul.....	49
Gambar 5. 1 IPA Matrix Ulasan Google Maps Review dan Instagram.....	58
Gambar 5. 2 Alur <i>Dashboard VoC</i>	66
Gambar 5. 3 <i>Dashboard Voice of Customer</i>	67

BAB I PENDAHULUAN

1.1 Latar Belakang

Sektor pariwisata merupakan salah satu pilar penting dalam pembangunan ekonomi Indonesia karena mampu menyumbang devisa sekaligus membuka lapangan pekerjaan. Menurut data resmi Badan Pusat Statistik (BPS) (2026), jumlah wisatawan nusantara di Indonesia menunjukkan peningkatan yang signifikan dari 524 juta perjalanan pada tahun 2020 menjadi 1.200 juta perjalanan pada tahun 2025. Kenaikan ini menunjukkan pemulihan dan pertumbuhan sektor pariwisata domestik secara bertahap setelah periode penurunan akibat pandemi, dengan lonjakan yang semakin tajam sejak tahun 2023. Peningkatan jumlah pergerakan wisatawan ini mencerminkan tingginya minat masyarakat terhadap destinasi wisata dalam negeri serta semakin aktifnya mobilitas domestik. Kondisi tersebut menuntut pengelola destinasi wisata untuk meningkatkan kualitas layanan dan pengelolaan berbasis data agar mampu memenuhi ekspektasi wisatawan yang terus berkembang.

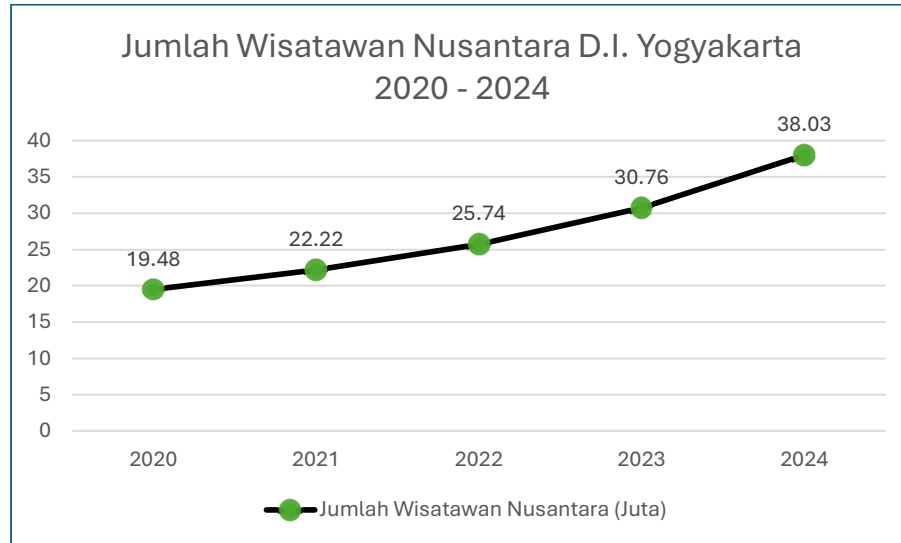


Gambar 1. 1 Perkembangan pariwisata di Indonesia

Sumber: (Badan Pusat Statistik, 2026)

Perkembangan sektor pariwisata juga tercermin di tingkat provinsi. Berdasarkan data Badan Pusat Statistik (BPS), jumlah wisatawan nusantara yang berkunjung ke Daerah Istimewa Yogyakarta menunjukkan tren peningkatan selama periode 2020–2024. Jumlah perjalanan wisatawan nusantara meningkat secara bertahap dari sekitar 19,48 juta perjalanan pada tahun

2020 menjadi 38,03 juta perjalanan pada tahun 2024. Kenaikan ini menunjukkan adanya pemulihan dan pertumbuhan aktivitas pariwisata domestik setelah periode pembatasan akibat pandemi COVID-19. Peningkatan jumlah wisatawan nusantara tersebut mengindikasikan bahwa Daerah Istimewa Yogyakarta tetap menjadi salah satu tujuan wisata domestik yang diminati di Indonesia.



Gambar 1. 2 Perkembangan pariwisata di D.I. Yogyakarta

Sumber: (Badan Pusat Statistik, 2026)

Namun, pertumbuhan pariwisata yang tinggi juga menimbulkan tantangan khususnya di destinasi wisata alam di Yogyakarta, salah satunya Goa Pindul di Kabupaten Gunungkidul. Wisata Goa Pindul memiliki jumlah pengunjung tahunan yang relatif stabil. Pada tahun 2019, jumlah pengunjung tahunannya adalah sebesar 100.000 pengunjung. Selanjutnya pada tahun 2023 jumlah pengunjung wisata Goa Pindul mengalami peningkatan menjadi 140.000 pengunjung. Wisata Goa Pindul menarik pengunjung karena pengalaman *cave tubing* yang unik, tetapi pertumbuhan kunjungan yang pesat berpotensi melampaui daya dukung lingkungan dan fasilitas (Siregar, 2025). Permasalahan lainnya yang ada di wisata Goa Pindul mencakup kualitas fasilitas umum, pelayanan, serta aksesibilitas yang dinilai masih perlu pengembangan lebih lanjut (Sukmaraga, 2016). Kondisi ini mencerminkan perlunya evaluasi persepsi wisatawan untuk mendukung pengelolaan destinasi wisata alam yang lebih efektif dan berkelanjutan.



Gambar 1. 3 Perkembangan jumlah pengunjung Goa Pindul tahunan

Sumber: (Siregar, 2025)

Analisis percakapan dan opini dalam industri pariwisata dapat dimanfaatkan untuk mengevaluasi persepsi pengunjung secara sistematis dengan teknik *text mining* seperti pemodelan topik dan analisis sentimen (Alamsyah et al., 2021). Dalam mengidentifikasi isu maupun persepsi pengunjung terhadap objek wisata seperti Goa Pindul, pendekatan analisis teks dan data menjadi penting. Salah satu metode yang relevan adalah *Latent Dirichlet Allocation* (LDA), yaitu teknik *topic modeling* yang digunakan untuk menemukan pola topik yang tersembunyi dalam kumpulan teks (Khatib Sulaiman et al., 2023). Metode ini dapat memetakan tema-tema utama diskusi dalam ulasan *online*, seperti komentar mengenai fasilitas, harga, keamanan, dan pengalaman wisatawan, sehingga lebih mudah untuk mengidentifikasi persepsi umum dan isu utama yang perlu ditindaklanjuti oleh pengelola destinasi wisata. Metode LDA secara luas dipakai dalam penelitian analisis sentimen untuk memahami opini publik dari data teks yang besar (Alamsyah et al., 2021).

Selain identifikasi topik, perlu pula mengevaluasi prioritas aspek apa saja yang harus diperbaiki berdasarkan persepsi pengunjung. Dalam konteks pengelolaan destinasi wisata, tidak seluruh aspek yang dibahas pengunjung memiliki tingkat kepentingan yang sama. Oleh karena itu, metode *Importance-Performance Analysis* (IPA) menjadi penting karena mampu memetakan atribut wisata berdasarkan tingkat kepentingan dan kinerja (Asyraini, 2023). Melalui pemetaan tersebut, pengelola dapat mengetahui aspek mana yang harus diprioritaskan untuk diperbaiki, dipertahankan, atau dievaluasi kembali secara lebih efisien. Penggunaan IPA dalam penelitian ini menjadi semakin relevan karena hasil analisis topik dan sentimen dari

ulasan digital belum cukup untuk menentukan strategi pengembangan tanpa adanya identifikasi prioritas perbaikan secara manajerial. Dengan demikian, integrasi IPA mampu menghasilkan rekomendasi pengembangan wisata Goa Pindul yang lebih tepat sasaran, berbasis data, dan sesuai dengan kebutuhan serta persepsi pengunjung.

Berdasarkan uraian di atas, penelitian ini diharapkan dapat memberi gambaran komprehensif tentang persepsi dan isu utama yang muncul dari ulasan wisatawan terhadap Goa Pindul melalui pemanfaatan metode LDA, analisis sentimen, dan IPA. Hasil penelitian ini diharapkan bermanfaat bagi pengambil kebijakan pariwisata lokal maupun pengelola destinasi dalam merumuskan strategi pengembangan wisata alam yang berkelanjutan dan berorientasi pada kepuasan pengunjung.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, peningkatan jumlah pengunjung wisata Goa Pindul belum diikuti dengan evaluasi yang sistematis terhadap persepsi pengunjung. Berbagai permasalahan seperti fasilitas, pelayanan, dan kenyamanan wisata masih perlu dikembangkan, namun belum diketahui aspek mana yang menjadi prioritas perbaikan berdasarkan sudut pandang wisatawan. Di sisi lain, ulasan pengunjung pada *Google Maps* dan Instagram yang mengandung informasi mengenai pengalaman wisata dapat dimanfaatkan secara optimal sebagai dasar evaluasi pengelolaan wisata. Oleh karena itu, berikut merupakan rumusan masalah dari penelitian ini:

1. Topik-topik utama apa saja yang terbentuk dari ulasan pengunjung terhadap Goa Pindul pada *Google Maps Review* dan Instagram berdasarkan metode *Latent Dirichlet Allocation* (LDA)?
2. Bagaimana kecenderungan sentimen ulasan pengunjung terhadap wisata Goa Pindul pada *Google Maps Review* dan Instagram?
3. Fasilitas atau atribut pelayanan apa saja yang menjadi prioritas perbaikan dalam pengelolaan wisata Goa Pindul berdasarkan hasil *Importance Performance Analysis* (IPA)?
4. Bagaimana rekomendasi strategi pengembangan wisata bagi pengelola Goa Pindul berdasarkan prioritas permasalahan yang didapatkan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah tertera, maka tujuan penelitian dapat ditentukan sebagai berikut:

1. Untuk mengidentifikasi topik-topik utama yang muncul dalam ulasan pengunjung terhadap wisata Goa Pindul pada *Google Maps Review* dan Instagram berdasarkan metode *Latent Dirichlet Allocation* (LDA).
2. Untuk mengetahui kecenderungan sentimen ulasan pengunjung terhadap wisata Goa Pindul pada *Google Maps Review* dan Instagram
3. Untuk menentukan prioritas perbaikan fasilitas dan atribut pelayanan wisata Goa Pindul berdasarkan hasil *Importance Performance Analysis* (IPA).
4. Menentukan rekomendasi strategi pengembangan wisata bagi pengelola Goa Pindul berdasarkan prioritas permasalahan yang didapatkan.

1.4 Manfaat Penelitian

Berikut merupakan manfaat dari adanya penelitian ini:

1. Bagi Peneliti
 - a. Untuk meningkatkan wawasan dan pengalaman peneliti terhadap penerapan metodologi penelitian Analisis Sentimen, *Latent Dirichlet Allocation* (LDA), dan *Importance Performance Analysis* (IPA).
 - b. Mengetahui penerapan keilmuan yang telah dipelajari dalam bidang Teknik Industri.
2. Bagi Universitas
 - a. Menjadi bukti relevansi akademik dan aplikatif dalam menjawab tantangan industri nasional.
 - b. Berkontribusi dalam meningkatkan reputasi universitas sebagai institusi yang menghasilkan penelitian inovatif serta relevan dengan kebutuhan industri dan masyarakat.
3. Bagi Industri dan Masyarakat
 - a. Menjadi dasar bagi industri pariwisata dalam melakukan pengembangan dan perbaikan fasilitas yang berorientasi terhadap tanggapan pengunjung.
 - b. Menjadi referensi untuk penelitian berikutnya dalam bidang pariwisata atau penelitian yang menggunakan metodologi Analisis Pemodelan, *Latent Dirichlet Allocation* (LDA), dan *Importance Performance Analysis* (IPA).

1.5 Batasan Penelitian

Berikut ini merupakan batasan penelitian:

1. Data yang digunakan bersumber dari *Google Maps Review* serta komentar instagram di akun @goapindul.id dan @guapindul_jogja pada periode Januari 2020 hingga Februari 2026, sehingga hasil dari penelitian ini bergantung pada kualitas data yang didapatkan dari sumber tersebut.
2. Metode LDA dan IPA dalam penelitian ini memiliki keterbatasan dalam menangkap konteks semantik yang kompleks serta bergantung pada interpretasi peneliti, sehingga hasil analisis belum sepenuhnya merepresentasikan makna kualitatif data secara menyeluruh.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Literatur

2.1.1. Penelitian Web Scrapping pada Pariwisata

Perkembangan platform *digital* seperti TripAdvisor, Google Maps, dan media sosial telah mendorong penggunaan teknik web scraping sebagai metode pengumpulan data dalam penelitian berbasis ulasan daring. Web scraping merupakan teknik ekstraksi data secara otomatis dari halaman web untuk memperoleh informasi dalam jumlah besar yang kemudian dapat diolah lebih lanjut menggunakan metode text mining dan analisis sentimen.

Dalam penelitian pariwisata, web scraping banyak digunakan untuk mengumpulkan data ulasan wisata secara sistematis. Ali et al. (2021) menggunakan teknik pengambilan data otomatis untuk mengekstraksi lebih dari 39.000 ulasan dari *platform* TripAdvisor yang berkaitan dengan berbagai *attraction* di Kota Marrakech. Data yang diperoleh melalui scraping tersebut kemudian dianalisis menggunakan pendekatan LDA dan analisis sentimen untuk mengidentifikasi topik dominan serta polaritas opini wisatawan. Hal ini menunjukkan bahwa web scraping berperan penting dalam menyediakan dataset berskala besar yang tidak memungkinkan diperoleh melalui survei manual.

Demikian pula, Ipmawati et al. (2024) melakukan proses *scraping* terhadap ulasan Google Maps pada berbagai destinasi wisata di Daerah Istimewa Yogyakarta menggunakan layanan Apify. Data yang terkumpul kemudian diproses melalui tahapan *preprocessing*, ekstraksi fitur TF-IDF, dan klasifikasi menggunakan *Support Vector Machine* (SVM). Penelitian ini menunjukkan bahwa *web scraping* memungkinkan pengumpulan data ulasan secara efisien dan real-time, sehingga mendukung analisis persepsi wisatawan berbasis data aktual.

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa web scraping merupakan komponen penting dalam penelitian berbasis ulasan digital, khususnya dalam konteks pariwisata. Teknik ini memungkinkan peneliti memperoleh data dalam jumlah besar secara otomatis, sehingga meningkatkan representativitas dan kekayaan informasi dibandingkan metode survei konvensional (Ali et al., 2021; Ipmawati et al., 2024). Namun demikian, penelitian terdahulu umumnya memanfaatkan *scraping* hanya untuk kebutuhan analisis sentimen atau *topic modeling* secara terpisah. Integrasi *web scraping* dengan pendekatan analisis yang lebih komprehensif, seperti penggabungan analisis sentimen, LDA, dan IPA dalam satu kerangka evaluasi manajerial masih terbatas.

Oleh karena itu, dalam penelitian ini *web scraping* digunakan sebagai tahap awal untuk mengumpulkan ulasan dari Google Maps dan Instagram secara sistematis, yang selanjutnya dianalisis menggunakan pendekatan terintegrasi guna menghasilkan rekomendasi prioritas perbaikan destinasi wisata.

2.1.2. Penelitian Analisis Sentimen dalam Pariwisata

Analisis sentimen telah banyak diterapkan dalam penelitian pariwisata untuk memahami persepsi wisatawan terhadap destinasi melalui ulasan daring. Beberapa penelitian menggunakan algoritma klasifikasi seperti *Naïve Bayes* dan *Support Vector Machine* untuk mengidentifikasi polaritas sentimen pada ulasan *Google Maps* maupun platform lainnya (Ipmawati et al., 2024; Sudiatmika et al., 2025). Hasil penelitian tersebut menunjukkan bahwa analisis sentimen mampu mengelompokkan opini pengunjung ke dalam kategori positif, negatif, dan netral serta mengidentifikasi kecenderungan kepuasan wisatawan terhadap suatu destinasi.

Selain itu, pendekatan *aspect-based sentiment analysis* juga digunakan untuk mengidentifikasi sentimen terhadap atribut tertentu seperti fasilitas, aksesibilitas, dan pelayanan (Al Farizal & Nugroho, 2025; Arianto & Budi, 2020). Pendekatan ini memberikan hasil yang lebih spesifik dibandingkan klasifikasi polaritas umum. Namun demikian, sebagian besar penelitian analisis sentimen masih terbatas pada pengukuran kecenderungan opini tanpa menggali struktur pembahasan yang mendasari sentimen tersebut. Dengan kata lain, metode ini mampu menunjukkan “arah opini”, tetapi belum sepenuhnya menjelaskan tema utama yang membentuk opini tersebut.

Penelitian lain juga menunjukkan bahwa pendekatan klasifikasi sentimen berbasis TF-IDF dan *Naïve Bayes* efektif dalam mengidentifikasi kecenderungan opini wisatawan pada destinasi berbasis *Google Maps* (Karo Karo et al., 2025). Selain itu, studi komparatif terhadap berbagai metode analisis sentimen menunjukkan bahwa pendekatan berbasis leksikon dan *machine learning* memiliki karakteristik performa yang berbeda tergantung pada struktur teks dan domain penelitian.

2.1.3. Penelitian Topic Modelling Menggunakan Latent Dirichlet Allocation (LDA)

Pendekatan *topic modeling* menggunakan *Latent Dirichlet Allocation* (LDA) banyak digunakan untuk mengidentifikasi topik utama dalam kumpulan ulasan wisata. Metode ini mampu mengelompokkan kata ke dalam beberapa topik berdasarkan distribusi probabilitas sehingga

efektif dalam mengungkap tema tersembunyi dalam data teks berskala besar (Ali et al., 2021; Habets et al., 2021).

Penelitian terdahulu menunjukkan bahwa LDA dapat mengidentifikasi aspek-aspek yang sering dibahas wisatawan seperti fasilitas, harga, keamanan, dan pengalaman pelayanan (Ariansyah & Indahyanti, 2024; Hapsari & Nataliani, 2024). Namun, penelitian berbasis LDA umumnya hanya berhenti pada tahap eksplorasi topik tanpa mengaitkan hasil tersebut dengan evaluasi tingkat kepuasan atau prioritas perbaikan layanan. Dengan demikian, LDA menjawab pertanyaan “apa yang dibahas”, tetapi belum menjawab “seberapa penting dan bagaimana kinerjanya”.

Studi integratif yang menggabungkan LDA dengan analisis sentimen juga menunjukkan bahwa pendekatan ini mampu mengidentifikasi topik dominan sekaligus polaritas opini terhadap setiap topik yang terbentuk (Ali Laudza et al., 2025). Namun, penelitian tersebut belum mengaitkan hasil analisis dengan prioritas perbaikan berbasis evaluasi manajerial.

2.1.4. Penelitian Importance Analysis (IPA) dalam Evaluasi Destinasi

Importance Performance Analysis (IPA) merupakan metode yang banyak digunakan dalam penelitian pariwisata untuk menentukan prioritas perbaikan atribut layanan berdasarkan tingkat kepentingan dan kinerja. Metode ini diterapkan melalui pemetaan atribut ke dalam empat kuadran untuk mengidentifikasi aspek yang perlu diprioritaskan, dipertahankan, atau dikurangi (Lasarudin & Yoanda, 2024; Suprihatin & Rofiq, 2023).

Penelitian lainnya menunjukkan bahwa IPA efektif dalam memberikan rekomendasi manajerial berbasis persepsi pengunjung terhadap kualitas layanan destinasi wisata (Bratić et al., 2024; Rahman et al., 2025). Namun demikian, sebagian besar penelitian IPA masih menggunakan pendekatan survei kuesioner konvensional sebagai sumber data utama. Hal ini menyebabkan keterbatasan pada jumlah responden dan kurang memanfaatkan data ulasan digital yang tersedia secara luas dan *real-time*.

2.1.5. Research Gap dan Posisi Penelitian

Berdasarkan telaah penelitian terdahulu, masing-masing metode memiliki keunggulan dan keterbatasan. Analisis sentimen mampu mengidentifikasi kecenderungan opini wisatawan (Ipmawati et al., 2024; Sudiatmika et al., 2025), LDA mampu memetakan topik utama dalam

ulasan wisata (Ali et al., 2021; Habets et al., 2021), dan IPA mampu menentukan prioritas perbaikan atribut layanan (Rahman et al., 2025; Suprihatin & Rofiq, 2023).

Namun, penelitian yang mengintegrasikan ketiga metode tersebut dalam satu kerangka analisis masih terbatas. Selain itu, sebagian besar penelitian hanya memanfaatkan satu *platform* ulasan seperti Google Maps atau TripAdvisor, sementara integrasi lintas *platform* seperti Google Maps dan Instagram masih jarang dilakukan (Blanco-Moreno et al., 2024).

Oleh karena itu, penelitian ini mengintegrasikan analisis sentimen, pemodelan topik menggunakan LDA, serta *Importance Performance Analysis* (IPA) untuk menganalisis ulasan wisata Goa Pindul yang bersumber dari Google Maps dan Instagram. Pendekatan ini diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai persepsi pengunjung sekaligus menghasilkan rekomendasi prioritas perbaikan berbasis data digital.

Tabel 2. 1 Posisi penelitian

No	Penulis	Sumber Data			Metode			Objek Penelitian
		Tradisional (wawancara, kuesioner, observasi langsung)	Online Review Platform	Media Sosial	Analisis Sentimen	Latent Dirichlet Allocation (LDA)	Importance- Performance Analysis (IPA)	
1	(Ali et al., 2021)		✓		✓	✓		Destinasi wisata Kota Marrakech (TripAdvisor)
2	(Ipmawati et al., 2024)		✓		✓			Destinasi wisata di DIY (Google Maps)
3	(Habets et al., 2021)	✓						LinkedIn
4	(Sudiatmika et al., 2025)		✓		✓			Ulasan <i>Google Maps</i> Pura Besakih
5	(Ariansyah & Indahyanti, 2024)					✓		Korpus Teks Umum

No	Penulis	Sumber Data			Metode			Objek Penelitian
		Tradisional (wawancara, kuesioner, observasi langsung)	Online Review Platform	Media Sosial	Analisis Sentimen	Latent Dirichlet Allocation (LDA)	Importance- Performance Analysis (IPA)	
6	(Bratić et al., 2024)	✓					✓	Nisville Jazz Festival
7	(Blanco-Moreno et al., 2024)			✓				Konten destinasi wisata (Instagram)
8	(Arianto & Budi, 2020)		✓		✓			Ulasan Borobudur & Prambanan (Google Maps)
9	(Hapsari & Nataliani, 2024)			✓		✓		Komentar aplikasi/teks
10	(Karo Karo et al., 2025)		✓		✓			Waterfront City (Google Maps)
11	(Al Farizal & Nugroho, 2025)		✓		✓			Ulasan destinasi wisata (Google Maps)

No	Penulis	Sumber Data			Metode			Objek Penelitian
		Tradisional (wawancara, kuesioner, observasi langsung)	Online Review Platform	Media Sosial	Analisis Sentimen	Latent Dirichlet Allocation (LDA)	Importance- Performance Analysis (IPA)	
12	(Suprihatin & Rofiq, 2023)	✓					✓	Wisata bahari
13	(Lasarudin & Yoanda, 2024)	✓					✓	Layanan jasa <i>travel</i>
14	(Rahman et al., 2025)	✓					✓	Taman Nasional Komodo
15	(Ali Laudza et al., 2025)		✓	✓	✓	✓		Ulasan wisata <i>multi-platform</i>
16	Penelitian Ini		✓	✓	✓	✓	✓	Ulasan Wisata Goa Pindul <i>(Google Maps dan Instagram)</i>

2.2 Landasan Teori

2.2.1. Pariwisata

Istilah pariwisata berasal dari bahasa Sanskerta yang tersusun dari dua kata, yaitu *pari* dan *wisata*. Kata *pari* bermakna banyak atau berkeliling, sedangkan *wisata* berarti pergi atau bepergian (Wirawan & Semara, 2021). Sedangkan menurut Fauzi et al. (2022), berdasarkan Undang-Undang Republik Indonesia Nomor 9 Tahun 1990, pariwisata diartikan sebagai segala sesuatu yang berkaitan dengan kegiatan perjalanan wisata, termasuk di dalamnya perusahaan, daya tarik dan atraksi wisata, serta berbagai usaha yang mendukung penyelenggaraan kegiatan pariwisata. Kontribusi pariwisata nasional semakin signifikan bagi Indonesia karena sektor ini memberikan dampak besar terhadap perekonomian. Pariwisata mampu menghasilkan devisa yang cukup besar untuk mendukung pengadaan barang modal yang diperlukan dalam kegiatan produksi. Selain itu, pengembangan pariwisata turut memacu pembangunan infrastruktur. Sektor ini juga berperan dalam mendorong pertumbuhan berbagai sektor ekonomi lainnya (Satrio Wibowo & Arviana Belia, 2023).

2.2.2. Data Scrapping

Data *scrapping* atau *web scrapping* merupakan teknik pengambilan data terstruktur dari situs web secara otomatis menggunakan aplikasi atau skrip pemrograman tertentu, di mana data yang diperoleh dapat mencakup jumlah yang sangat besar, mulai dari ribuan hingga miliaran entri yang tersebar di internet (Rudini et al., 2023). Tahap persiapan data dalam penelitian ini melibatkan pemanfaatan data media sosial yang bersifat sangat dinamis karena mencerminkan perilaku manusia serta kondisi sosial yang sedang berlangsung di masyarakat. Namun, data tersebut belum memberikan nilai guna apabila tidak melalui proses pengolahan terlebih dahulu. Oleh karena itu, diperlukan aktivitas ekstraksi data dari web atau media sosial untuk mengungkap pola-pola tersembunyi di dalamnya, yang dikenal sebagai proses *scrapping* (Yuniar et al., 2022).

2.2.3. Python

Python merupakan bahasa pemrograman yang diciptakan oleh Guido Van Rossum pada awal tahun 1990 untuk keperluan umum di Centrum voor Wiskunde and Informatica (CWI) (Baskara, 2022). Bahasa pemrograman Python mengintegrasikan fleksibilitas tingkat tinggi,

kemudahan dibaca, serta antarmuka yang terstruktur dengan kapabilitas tingkat rendah melalui dukungan antarmuka C resmi, sehingga memungkinkan perluasan bahasa menggunakan kode C dan integrasi dengan pustaka pihak ketiga seperti C, C++, dan Fortran. Kemampuan ini menjadikan Python sangat sesuai untuk komputasi ilmiah modern karena menyediakan keseimbangan antara performa dan kemudahan pengembangan, sekaligus memberi keleluasaan dalam memilih fitur tingkat rendah maupun tinggi sesuai kebutuhan (Pérez et al., 2011).

2.2.4. *Google Maps Review*

Google Maps merupakan salah satu platform berbasis web yang memungkinkan pengguna untuk memberikan ulasan terhadap lokasi yang telah dikunjungi. Melalui *Google Maps*, dapat diperoleh data berupa *review* masyarakat yang berisi komentar dan penilaian terhadap objek wisata, khususnya destinasi wisata Goa Pindul. Ulasan tersebut mencerminkan persepsi pengunjung, baik dalam bentuk tanggapan positif maupun negatif. Untuk mengidentifikasi kecenderungan sentimen dari pendapat masyarakat tersebut, dapat diterapkan metode analisis sentimen guna mengklasifikasikan opini pengguna berdasarkan polaritasnya (Khofifah et al., 2022).

2.2.5. *Instagram*

Instagram adalah salah satu *platform* media sosial yang sangat populer di dunia, khususnya di Indonesia. Menurut laporan dari We Are Social dan Meltwater, jumlah penggunanya telah menembus angka 1,91 miliar per Oktober 2025. Berdasarkan hal tersebut, ini berarti hampir seperempat (23,1%) dari seluruh populasi global kini aktif menggunakan platform berbagi foto tersebut. Dari jumlah data tersebut, Indonesia berada diposisi keempat dengan jumlah 107,6 juta pengguna aktif (Pau, 2025). Instagram memungkinkan pengguna menampilkan foto secara instan melalui platformnya. Istilah “gram” sendiri berasal dari kata “telegram”, yang merujuk pada sistem penyampaian informasi secara cepat. Sejalan dengan konsep tersebut, Instagram memfasilitasi pengunggahan foto melalui jaringan internet sehingga informasi visual dapat diterima oleh pengguna lain dalam waktu singkat (Damiana, 2025).

2.2.6. *Text Mining*

Pada era saat ini, penerapan text mining menjadi sangat penting untuk memvisualisasikan serta mengevaluasi pengetahuan yang diperoleh dari kumpulan dokumen teks dalam jumlah besar

(Deolika & Taufiq Luthfi, 2019). Penerapan text mining dalam suatu studi kasus perlu mengikuti tahapan analisis yang sistematis. Oleh karena itu, sebelum data teks diproses menggunakan berbagai metode text mining, tahap awal yang harus dilakukan adalah text preprocessing. Berikut ini merupakan tahapan *text preprocessing* pada *text mining* (Adhe et al., 2020):

1. *Case Folding*

Case folding merupakan tahapan normalisasi teks dengan mengonversi seluruh karakter menjadi huruf kecil guna menyeragamkan bentuk kata.

2. *Filtering*

Filtering merupakan tahap seleksi token untuk memperoleh kata-kata bermakna yang merepresentasikan isi dokumen, dilakukan dengan menghapus stopword serta mengubah kata berimbuhan ke bentuk dasarnya; stopword sendiri merupakan kata umum yang tidak memiliki kontribusi signifikan terhadap makna teks.

3. *Tokenizing*

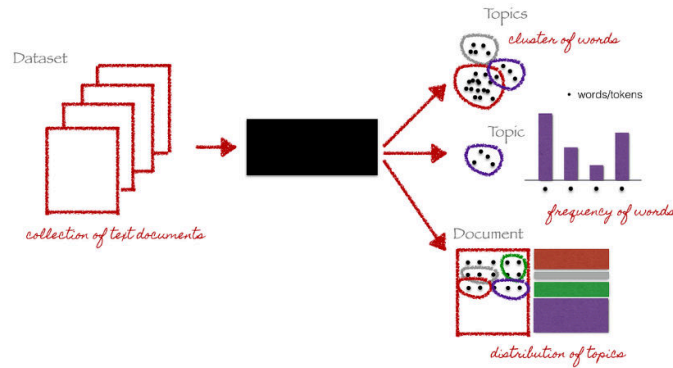
Tokenizing adalah proses pemisahan teks menjadi unit kata, sekaligus menghilangkan angka, tanda baca, serta karakter non-alfabet agar hanya tersisa kata yang relevan.

4. *Stemming*

Stemming adalah proses pengembalian kata ke bentuk dasarnya setelah tahap filtering, di mana penelitian ini menggunakan algoritma Porter dari library NLTK Python.

2.2.7. *Latent Dirichlet Allocation* (LDA)

Latent Dirichlet Allocation (LDA) adalah metode *text mining* yang digunakan untuk menemukan topik-topik tersembunyi dalam kumpulan dokumen. LDA bekerja dengan mengidentifikasi pola kemunculan kata sehingga setiap dokumen direpresentasikan sebagai kombinasi beberapa topik dengan proporsi tertentu. (Khatib Sulaiman et al., 2023). Berikut merupakan alur kerja dari LDA:



Gambar 2. 1 Alur kerja metode LDA

Tahapan kerja *Latent Dirichlet Allocation* (LDA) diawali dengan inisialisasi sejumlah parameter, seperti jumlah dokumen, jumlah topik, serta jumlah iterasi, di mana penentuan jumlah topik (k) menjadi parameter utama. Selanjutnya, setiap kata dialokasikan secara acak ke dalam topik tertentu berdasarkan distribusi Dirichlet. Proses ini kemudian diulang secara iteratif untuk seluruh kata dalam korpus hingga model mencapai kondisi konvergen (Nurlayli & Nasichuddin, 2019).

Metode *Latent Dirichlet Allocation* (LDA) secara matematis dapat dimodelkan sebagai berikut (Purnama et al., 2025):

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^N p(z_n|\theta) p(w_n|z_n, \beta) \quad (2.1)$$

Keterangan:

- $p(\theta, z, w|\alpha, \beta)$: Probabilitas gabungan
- θ : Disribusi topik dalam satu dokumen
- Z_n : Topik yang ditugaskan untuk kata ke- n
- w_n : Kata ke- n dalam dokumen
- α : Hyperparameter Dirichlet untuk distribusi topik (θ).
- β : Hyperparameter distribusi kata dalam topik.

Tahap berikutnya adalah menentukan jumlah topik yang paling optimal. Proses ini dilakukan dengan menghitung nilai koherensi untuk setiap variasi jumlah topik yang diuji, kemudian membandingkan hasilnya guna memperoleh jumlah topik yang paling representatif. Perhitungan nilai koherensi tersebut dilakukan menggunakan persamaan berikut.

$$Coherence(t; V^{(t)}) = \sum_{m=2}^M \sum_{l=1}^{m-1} \log \frac{D(V_m^{(t)}, V_l^{(t)}) + 1}{D(V_{ml}^{(t)})} \quad (2.2)$$

Keterangan:

- t : Indeks topik ke- t , menunjukan topik mana yang sedang dibahas
- $V^{(t)}$: Himpunan kata-kata utama pada topik ke- t
- $D(v)$: *Document frequency* dari kata v
- $D(V_m^{(t)}, V_l^{(t)})$: Frekuensi kemunculan kata v_m dan v_l bersama.

Jumlah topik optimal ditentukan dengan menghitung *coherence score* pada setiap variasi jumlah topik, kemudian dipilih nilai yang paling representatif. Dalam metode LDA, dua keluaran utama yang diperoleh adalah pengelompokan topik dan tingkat kepentingan masing-masing topik. Tingkat kepentingan topik dihitung dengan menjumlahkan bobot setiap kata kunci dalam topik, kemudian dinormalisasi untuk memperoleh *importance score* (Purnama et al., 2023). Perhitungan *importance score* dapat ditulis dalam persamaan berikut:

$$WI_t = \sum_{i=0}^n Weight_i \quad (2.3)$$

$$Importance\ score = \frac{WI_t - WI_{t_{min}}}{WI_{t_{max}} - WI_{t_{min}}} \quad (2.4)$$

Keterangan:

- WI_t = total bobot kepentingan topik ke- t
- $Weight_i$ = bobot kata kunci ke- i dalam topik
- WI_{max} = nilai bobot topik terbesar
- WI_{min} = nilai bobot topik terkecil
- n = jumlah kata kunci dalam topik

Pada penelitian ini, *Latent Dirichlet Allocation* dipilih sebagai metode pemodelan topik berdasarkan beberapa kriteria yang sesuai dengan data pada penelitian ini. Berikut ini merupakan tabel perbandingan metode pemodelan topik antara *Latent Dirichlet Allocation* (LDA), *Non-Negative Matrix Factorization* (NMF), *Topic to Vector* (Top2Vec), dan *Bidirectional Encoder Representations from Transformers Topic* (BERTopic) (Egger & Yu, 2022):

Tabel 2. 2 Perbandingan metode pemodelan topik

Metode	Tidak Perlu Domain Knowledge	Multi-Topik per Dokumen	Tentukan Jumlah Topik	Komputasi Efisien	Struktur Probabilistik
LDA	✓	✓	✓	✓	✓
NMF	✓	✓	✓	✓	X
Top2Vec	✓	X	X	✓	X
BERTopic	✓	X	X	X	X

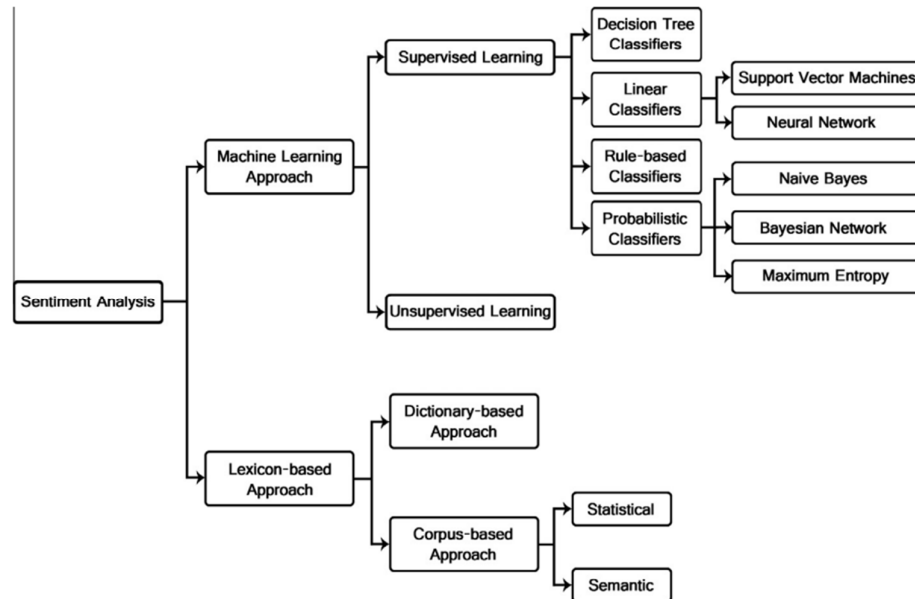
Berdasarkan tabel perbandingan metode pemodelan tersebut, keempat metode memiliki karakter yang sama, yaitu tidak memerlukan *domain knowledge*. Metode LDA dan NMF memiliki karakter dapat melakukan multi topik per dokumen, menentukan jumlah topik, dan komputasi yang efisien. Akan tetapi metode LDA memiliki keunggulan dalam memiliki struktur yang probabilistik. Dalam hal itu, metode LDA lebih fleksibel dalam penentuan jumlah topik. Sedangkan karakter dari metode Top2Vec dan BERTopic tidak dapat melakukan multi topik per dokumen, jumlah topik yang hanya satu per dokumen, dan tidak memiliki struktur probabilistik. Selain itu, metode BERTopic memiliki karakter yang berat dalam komputasi. Oleh sebab itu, LDA dipilih karena mampu merepresentasikan dokumen dalam bentuk distribusi probabilistik, menghasilkan topik yang stabil, serta memiliki efisiensi komputasi yang baik dan umum digunakan dalam penelitian.

2.2.8. Analisis Sentimen

Analisis sentimen adalah teknik dalam text mining yang berfokus pada identifikasi opini atau perasaan yang tersirat dalam teks. Teknik ini telah banyak diterapkan untuk memahami persepsi pengguna terhadap suatu produk atau layanan berdasarkan ulasan daring (*online reviews*). Menurut Medhat, Hassan, & Korashy (2014), analisis sentimen efektif untuk mengevaluasi opini dalam berbagai domain teks termasuk opini pengguna di media sosial dan *platform review*.

Dalam konteks pariwisata, analisis sentimen berperan penting untuk menilai pengalaman wisatawan terhadap destinasi berdasarkan ulasan mereka di platform digital seperti *Google Maps* dan media sosial. Data ulasan tersebut dapat mengungkapkan sentimen pengunjung terhadap aspek layanan, fasilitas, dan pengalaman secara keseluruhan (Kim et al., 2020).

Dengan mengolah ulasan ini, peneliti dapat menentukan kecenderungan sentimen serta faktor-faktor yang mempengaruhi kepuasan wisatawan.



Gambar 2. 2 Pengelompokan metode analisis sentimen

Sumber: (Medhat et al., 2014)

Pendekatan analisis sentimen dapat dikelompokkan menjadi dua, yaitu pendekatan berbasis *machine learning* dan berbasis leksikon. Pendekatan berbasis *machine learning* menggunakan model klasifikasi yang dilatih pada data berlabel, sedangkan pendekatan berbasis leksikon memanfaatkan kamus kata dengan skor polaritas tertentu untuk menilai sentimen teks (Medhat et al., 2014). Pendekatan berbasis leksikon sering dipilih untuk dataset ulasan singkat dan informal karena tidak memerlukan pelabelan data manual.

2.2.9. *Valence Aware Dictionary and sEntiment Reasoner (VADER)*

VADER (*Valence Aware Dictionary and sEntiment Reasoner*) merupakan salah satu metode analisis sentimen berbasis leksikon yang dirancang untuk menangani karakteristik bahasa informal dan singkat yang umum terdapat pada media sosial serta platform ulasan online. Metode ini memanfaatkan kamus kata dengan bobot valensi emosional dan aturan heuristik untuk menangkap intensitas emosional dalam teks, termasuk penggunaan huruf kapital, tanda baca, dan kata penguat (*intensifiers*).

Dalam literatur analisis sentimen general, pendekatan berbasis leksikon seperti VADER menunjukkan performa kompetitif saat menangani teks berformat pendek dan ekspresif, seperti komentar media sosial atau review singkat (Hutto & Gilbert, 2014). Pendekatan ini cocok untuk penelitian yang menggunakan ulasan media sosial seperti komentar Instagram dan review *Google Maps*, karena teks dalam format tersebut sering tidak mengikuti struktur formal bahasa.

Metode berbasis leksikon menghasilkan skor polaritas positif, negatif, dan netral, serta skor agregat yang dikenal sebagai *compound score* yang dinormalisasi dalam rentang -1 sampai $+1$ untuk menunjukkan arah dan kekuatan sentimen secara keseluruhan (Hutto & Gilbert, 2014). Dengan karakteristik tersebut, pendekatan berbasis leksikon seperti VADER menjadi pilihan yang relevan dan efisien dalam penelitian analisis sentimen berbasis teks media sosial dan ulasan *online*.

Pada penelitian ini, analisis sentimen dilakukan dengan pendekatan berbasis leksikon seperti VADER dengan beberapa pertimbangan karakter yang dimilikinya. Berikut merupakan tabel perbandingan metode analisis sentimen (Mehenaoui et al., 2024):

Tabel 2. 3 Perbandingan metode analisis sentimen

Metode	Tanpa Training Data	Berbasis Leksikon	Cocok Teks Sosial Media	Cepat & Efisien	Bisa Tangani Emoji/Slang
VADER	✓	✓	✓	✓	✓
TextBlob	✓	✓	✗	✓	✗
SentiWordNet	✓	✓	✗	✗	✗
SpaCy	✗	✗	✗	✓	✗

Pada tabel di atas, metode VADER memiliki karakteristik seperti tidak memerlukan *training data*, berbasis leksikon, cocok untuk ulasan teks sosial media, cepat & efisien, serta dapat menangani emoji atau teks *slang*. Oleh sebab itu, metode VADER digunakan dalam penelitian ini untuk analisis sentimen karena dinilai memiliki karakter yang paling sesuai dengan data dari penelitian ini yang bersumber dari sosial media dan memiliki emoji pada beberapa ulasan.

2.2.10. Importance-Performance Analysis (IPA)

Importance-Performance Analysis (IPA) pertama kali dikemukakan oleh Martilla dan James pada tahun 1977 sebagai suatu metode untuk mengukur tingkat kepuasan pelanggan terhadap produk maupun jasa. Pendekatan IPA memandang kepuasan sebagai fungsi dari dua komponen utama, yaitu kualitas produk atau layanan yang diterima oleh pelanggan serta kinerja perusahaan dalam memberikan produk atau jasa tersebut. Dengan demikian, IPA tidak hanya menganalisis aspek kinerja, tetapi juga memperhatikan kualitas sebagai faktor penentu dalam membentuk kepuasan responden (Asyraini, 2023).

Dalam penerapannya, *Importance-Performance Analysis* (IPA) memanfaatkan sebuah matriks yang dikenal sebagai Diagram Kartesius. Pada diagram tersebut, tingkat kinerja (*performance*) direpresentasikan pada sumbu horizontal (X), sedangkan tingkat kepentingan (*importance*) ditunjukkan pada sumbu vertikal (Y) (Pratiwi, 2018).

I M P O R T A N C E	Quadrant I Concentrate Here High Importance Low Performance	Quadrant II Keep Up the Good Work High Importance High Performance
	Quadrant III Low Priority Low Importance Low Performance	Quadrant IV Possible Overkill Low Importance High Performance

PERFORMANCE

Gambar 2. 3 IPA matrix

Sumber: (Murdy, 2012)

Penjelasan dari Gambar 2.3 adalah sebagai berikut:

1. Kuadran I (*Concentrate Here*)

Kuadran I memuat atribut yang dinilai sangat penting oleh pelanggan, namun tingkat kinerjanya belum memenuhi harapan sehingga kepuasan masih tergolong rendah. Oleh karena itu, variabel-variabel yang berada pada kuadran ini perlu menjadi fokus utama perbaikan guna meningkatkan kepuasan pelanggan.

2. Kuadran B (*Keep Up the Good Work*)

Kuadran II terdiri atas atribut yang dianggap penting oleh pelanggan dan telah menunjukkan kinerja yang sesuai dengan harapan, sehingga tingkat kepuasannya relatif

tinggi. Variabel pada kuadran ini perlu dipertahankan karena mencerminkan keunggulan produk atau jasa di mata pelanggan.

3. Kuadran III (*Low Priority*)

Kuadran III mencakup atribut yang dinilai kurang penting oleh pelanggan serta memiliki kinerja yang tidak menonjol. Variabel-variabel dalam kuadran ini dapat dipertimbangkan kembali, mengingat kontribusinya terhadap manfaat yang dirasakan pelanggan relatif kecil.

4. Kuadran IV (*Possible Overkill*)

Kuadran IV berisi atribut yang dianggap kurang penting oleh pelanggan, namun memiliki tingkat kinerja yang berlebihan. Oleh sebab itu, variabel-variabel pada kuadran ini dapat dikurangi atau dialokasikan ulang agar perusahaan dapat melakukan efisiensi biaya.

Rumus yang digunakan adalah sebagai berikut:

$$\bar{X} = \frac{\sum X_i}{n} \quad (2.5)$$

$$\bar{Y} = \frac{\sum Y_i}{n} \quad (2.6)$$

Keterangan:

- $\bar{X}$ = skor rata-rata tingkat kerja
- $\bar{Y}$ = skor rata-rata tingkat kepentingan
- $\sum X_i$ = jumlah total skor kinerja
- $\sum Y_i$ = jumlah total skor kepentingan
- n = jumlah responden

Selanjutnya dihitung $\bar{\bar{X}}$ (rata-rata dari rata-rata skor tingkat kinerja) dan $\bar{\bar{Y}}$ (rata-rata dari rata-rata skor tingkat kepentingan). Rumus yang digunakan adalah sebagai berikut (Pratiwi, 2018):

$$\bar{\bar{X}} = \frac{\sum_{i=1}^n \bar{X}_i}{k} \quad (2.7)$$

$$\bar{\bar{Y}} = \frac{\sum_{i=1}^n \bar{Y}_i}{k} \quad (2.8)$$

Keterangan:

- K = banyaknya atribut pelayanan dalam kuesioner yang diberikan kepada responden

Metode IPA memiliki beberapa keunggulan dibandingkan metode lainnya. Keunggulan tersebut meliputi kemampuannya dalam mengidentifikasi atribut produk atau jasa yang perlu diperbaiki maupun dikurangi guna mempertahankan tingkat kepuasan pelanggan. Selain itu, hasil analisis relatif mudah dipahami, skala pengukurannya sederhana, serta penerapannya tidak memerlukan biaya yang besar (Dian Anggraini et al., 2015).

BAB III

METODOLOGI PENELITIAN

3.1 Subjek Penelitian

Subjek penelitian dalam penelitian ini adalah wisatawan atau pengunjung yang pernah berkunjung ke wisata Goa Pindul dan memberikan ulasan atau komentar secara daring melalui *platform Google Maps Review* dan Instagram. Ulasan dan komentar tersebut digunakan sebagai representasi persepsi, opini, dan pengalaman pengunjung terhadap wisata Goa Pindul.

3.2 Objek Penelitian

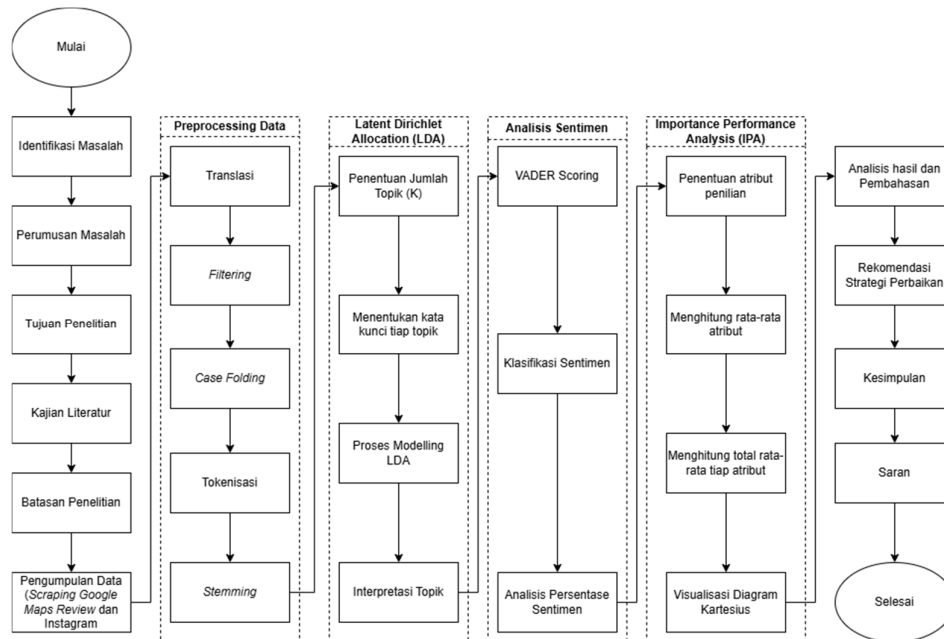
Objek penelitian dalam penelitian ini adalah persepsi dan sentimen pengunjung terhadap wisata Goa Pindul yang tercermin dalam ulasan dan komentar pada platform *Google Maps Review* dan Instagram, yang dianalisis untuk mengidentifikasi topik utama pembahasan serta menentukan prioritas perbaikan dalam pengelolaan wisata menggunakan pendekatan analisis sentimen, metode *Latent Dirichlet Allocation* (LDA), dan *Importance Performance Analysis* (IPA).

3.3 Sumber dan Jenis Data

Sumber dan jenis data dalam penelitian ini adalah data sekunder yang diperoleh dari ulasan dan komentar pengunjung wisata Goa Pindul pada platform *Google Maps Review* dan Instagram pada akun @goapindul.id dan @guapindul_jogja. Data tersebut berupa teks opini dan pengalaman pengunjung yang dikumpulkan melalui teknik *web scraping* menggunakan bahasa pemrograman Python. Data teks ini dianalisis menggunakan tahapan *text preprocessing*, analisis sentimen, dan pemodelan topik dengan metode *Latent Dirichlet Allocation* (LDA). Selain itu, penelitian ini juga menghasilkan data kuantitatif berupa skor sentimen, distribusi topik, serta nilai rata-rata tingkat kepentingan dan kinerja atribut wisata yang digunakan dalam analisis *Importance Performance Analysis* (IPA) untuk menentukan prioritas perbaikan pengelolaan wisata.

3.4 Alur Penelitian

Berikut merupakan alur penelitian yang diterapkan pada penelitian ini:



Gambar 3. 1 Alur penelitian

Alur penelitian ini disusun secara sistematis mulai dari tahap perumusan masalah hingga penarikan kesimpulan. Berdasarkan diagram alir penelitian, tahapan penelitian dapat dijelaskan sebagai berikut:

1. Mulai
2. Identifikasi Masalah

Tahap awal penelitian dimulai dengan mengidentifikasi permasalahan yang terdapat pada objek penelitian, yaitu wisata Goa Pindul. Permasalahan yang diidentifikasi berkaitan dengan persepsi dan sentimen pengunjung terhadap destinasi wisata yang tercermin dalam ulasan daring pada platform *Google Maps* dan Instagram.

3. Perumusan Masalah

Setelah masalah diidentifikasi, dilakukan perumusan masalah secara spesifik untuk menentukan fokus penelitian, yaitu analisis sentimen, identifikasi topik utama menggunakan metode *Latent Dirichlet Allocation* (LDA), serta penentuan prioritas perbaikan menggunakan *Importance Performance Analysis* (IPA).

4. Tujuan Penelitian

Tahap ini bertujuan untuk menetapkan sasaran yang ingin dicapai, yaitu mengetahui kecenderungan sentimen pengunjung, mengidentifikasi topik utama yang terbentuk dari ulasan, serta menentukan prioritas perbaikan berdasarkan persepsi pengunjung.

5. Kajian Literatur

Kajian literatur dilakukan untuk memperkuat landasan teori yang digunakan dalam penelitian, meliputi teori analisis sentimen, *text mining*, *Latent Dirichlet Allocation* (LDA), dan *Importance Performance Analysis* (IPA).

6. Batasan Penelitian

Pada tahap ini ditetapkan ruang lingkup penelitian, termasuk sumber data, periode pengambilan data, serta metode analisis yang digunakan agar penelitian lebih terarah dan terfokus.

7. Pengumpulan Data

Data penelitian diperoleh melalui proses scraping ulasan pengunjung dari *platform Google Maps Review* dan Instagram yang berkaitan dengan wisata Goa Pindul pada periode Januari 2020 hingga Februari 2026. Data yang dikumpulkan berupa teks ulasan dan komentar yang bersifat publik.

8. *Preprocessing Data*

Tahap *preprocessing* dilakukan untuk membersihkan dan menyiapkan data teks sebelum dianalisis lebih lanjut. Tahapan *preprocessing* meliputi:

a. Translasi

Karena metode VADER menggunakan kamus Bahasa Inggris, ulasan yang berbahasa Indonesia diterjemahkan ke dalam Bahasa Inggris untuk meningkatkan akurasi analisis sentimen. Proses translasi ini menggunakan *software Google Docs* untuk akurasi dan efisiensi penerjemahan teks dalam jumlah yang besar.

b. *Filtering*

Tahap *filtering* bertujuan untuk menghapus kata-kata yang tidak memiliki makna penting dalam analisis (*stopwords*), seperti kata hubung dan kata umum lainnya.

c. *Case Folding*

Pada tahap ini seluruh teks diubah menjadi huruf kecil (*lowercase*) untuk menyeragamkan format penulisan dan menghindari perbedaan kata akibat penggunaan huruf kapital.

d. *Tokenizing*

Proses *tokenizing* dilakukan dengan memecah teks menjadi unit-unit kata (token) agar dapat dianalisis secara terpisah.

e. *Stemming*

Stemming dilakukan untuk mengubah kata berimbuhan menjadi bentuk dasar sehingga variasi kata dengan makna yang sama dapat dikelompokkan dalam satu bentuk.

9. *Latent Dirichlet Allocation (LDA)*

a. Penentuan jumlah topik (k)

Pada tahap ini, proses pemodelan topik dilakukan dengan memanfaatkan *library Python*, yang mendukung implementasi LDA secara fleksibel dan efisien. Model tersebut bekerja dengan mengelompokkan kata-kata yang memiliki kemunculan bersamaan dalam dokumen ke dalam satu topik tertentu, sehingga dapat menghasilkan representasi tematik dari keseluruhan kumpulan data. Jumlah topik ditentukan melalui pengujian beberapa nilai K dan evaluasi menggunakan *coherence score* seperti pada persamaan (2.2).

b. Menentukan kata kunci tiap topik

Setelah model dengan *conherence score* terbaik didapatkan, topik dengan kata-kata kunci dominan akan disatukan. Setiap topik diidentifikasi atau diberi label berdasarkan makna umum dari topik tersebut.

c. Proses *modelling* LDA

Dilakukan pemodelan probabilistik untuk memperoleh distribusi topik pada setiap dokumen. Pada tahap ini juga ditentukan nilai bobot *importance* pada tiap-tiap topik dengan menggunakan persamaan (2.3) dan (2.4). Nilai *Importance* diperoleh melalui proses agregasi bobot kepentingan dari setiap kata kunci yang terdapat dalam suatu topik. Setiap kata kunci memiliki nilai *importance* yang ditentukan berdasarkan frekuensi kemunculannya, serta nilai *weight* yang merepresentasikan probabilitas kata tersebut dalam topik sesuai dengan hasil pemodelan LDA.

d. Interpretasi topik

Topik yang terbentuk diinterpretasikan sesuai konteks wisata Goa Pindul.

10. Analisis sentimen

Tahapan analisis sentimen meliputi:

a. *VADER Scoring*

Setiap ulasan dianalisis menggunakan *Valence Aware Dictionary and sEntiment Reseanor (VADER)* untuk memperoleh skor sentimen berupa nilai positif, negatif,

netral, dan *compound*. Skor *compound* merupakan skor utama yang digunakan untuk mengklasifikasikan sentimen secara keseluruhan.

b. Klasifikasi Sentimen

Penentuan kategori sentimen dilakukan berdasarkan nilai *compound score*, di mana skor $\geq 0,05$ dikategorikan positif, $\leq -0,05$ negatif, dan di antara keduanya netral.

Tabel 3. 1 Klasifikasi sentimen

<i>Compound Score</i>	Klasifikasi Sentimen
$\geq 0,05$	Positif
$-0,05 < x < -0,05$	Netral
$\leq -0,05$	Negatif

c. Analisis Persentase Sentimen

Dilakukan perhitungan distribusi persentase sentimen untuk mengetahui kecenderungan opini pengunjung secara keseluruhan.

11. *Importance-Performance Analysis* (IPA)

a. Penentuan atribut penilaian

Atribut pelayanan ditentukan berdasarkan hasil interpretasi topik yang diperoleh dari pemodelan LDA dan analisis sentimen. Proses ini dilakukan untuk memastikan bahwa atribut yang dianalisis benar-benar merepresentasikan aspek pelayanan yang penting bagi pengunjung serta sesuai dengan konteks penelitian. Nilai atribut *importance* didapatkan dari nilai *importance weight* topik yang terbentuk, sedangkan nilai atribut *performance* didapatkan dari nilai *compound* topik tersebut.

b. Menghitung rata-rata atribut

Pada tahap ini dilakukan perhitungan rata-rata skor kepentingan (*importance*) dan kepuasan (*performance*) untuk setiap atribut yang telah ditetapkan. Nilai rata-rata tersebut digunakan sebagai dasar untuk menggambarkan persepsi umum responden terhadap masing-masing atribut pelayanan.

c. Menghitung total rata-rata atribut

Selanjutnya, dihitung nilai rata-rata keseluruhan dari seluruh atribut sebagai titik potong (*grand mean*) dalam Diagram Kartesius. Nilai ini berfungsi sebagai batas pembagi antara kategori tinggi dan rendah pada sumbu kepentingan dan kinerja.

d. Visualisasi diagram kartesius

Setiap atribut kemudian dipetakan ke dalam Diagram Kartesius berdasarkan nilai rata-rata kepentingan dan kepuasan yang telah diperoleh. Hasil pemetaan ini mengelompokkan atribut ke dalam empat kuadran, sehingga memudahkan dalam menentukan prioritas perbaikan dan strategi pengembangan layanan.

12. Analisis dan Pembahasan

Pada tahap ini dilakukan interpretasi terhadap hasil analisis sentimen, LDA, dan IPA untuk menjawab rumusan masalah penelitian. Hasil penelitian dibahas secara komprehensif dengan mengacu pada teori dan penelitian terdahulu.

13. Rekomendasi Strategi Perbaikan

Berisi rekomendasi strategi perbaikan bagi pengelola wisata Goa Pindul berdasarkan hasil prioritas permasalahan yang terbentuk dari IPA Matriks.

14. Kesimpulan

Tahap ini menyajikan jawaban atas rumusan masalah berdasarkan hasil analisis yang telah dilakukan.

15. Saran

Berisi rekomendasi serta usulan bagi peneliti selanjutnya.

16. Selesai

BAB IV

PENGOLAHAN DATA

4.1 Pengumpulan Data

4.1.1. *Data Scraping Google Maps*

Pada penelitian ini, pengumpulan data ulasan dari platform *Google Maps* dilakukan dengan metode *web scraping* menggunakan bahasa pemrograman *Python* dengan bantuan layanan pihak ketiga, yaitu *Apify Google Maps Reviews Scraper. Tools* tersebut digunakan untuk mengekstraksi data ulasan secara otomatis dari halaman *Google Maps Review* objek wisata Goa Pindul.

Proses *scraping* dilakukan pada rentang waktu Januari 2020 hingga Februari 2026. Berdasarkan hasil *scraping*, diperoleh sebanyak 5.137 data ulasan. Namun, tidak seluruh data memiliki teks komentar yang dapat dianalisis. Oleh karena itu, dilakukan proses seleksi data dengan hanya mengambil ulasan yang mengandung teks, sehingga diperoleh sebanyak 2.661 ulasan yang digunakan dalam penelitian ini.

4.1.2. *Data Scraping Instagram*

Pada penelitian ini, pengumpulan data dilakukan dengan metode *web scraping* menggunakan bahasa pemrograman *Python*. Proses *scraping* dilakukan dengan memanfaatkan layanan pihak ketiga, yaitu *Apify*, yang memungkinkan pengambilan data komentar dari media sosial Instagram secara otomatis. Data yang diambil berupa komentar pengguna pada beberapa unggahan (post dan reel) yang berkaitan dengan objek wisata Goa Pindul pada rentang waktu Januari 2020 hingga Februari 2026. Jumlah data yang didapatkan pada kedua akun *Instagram* @goapindul.id dan @guapindul_jogja sebanyak 861 komentar.

4.2 Preprocessing Data

4.2.1. *Translasi*

Tahap *preprocessing data* yang pertama dilakukan adalah tahap translasi. Pada tahap ini keseluruhan data hasil *scraping* yang telah didapatkan dilakukan translasi kedalam bahasa Inggris. Tahap translasi ini menggunakan *software* dari *google spreadsheet* dengan *function* “=GOOGLETRANSLATE(A1;"auto";"en")”, formula ini dapat melakukan proses translasi dari bahasa asal ke bahasa Inggris dengan jumlah *text* yang banyak. Berikut merupakan contoh *text* yang dilakukan translasi:

Tabel 4. 1 Hasil *text* translasi

No	Text	Text Translate
1	Ke sini waktu bulan puasa, jadi agak sepi dan tidak pakai antri. Pemandunya ahli banget, ramah, dan suka cerita, jadi tau ttg ceritanya. Kalau mau bawa hape, pastiin dipakein kantong dan bs dikalungin, karena sepanjang perjalanan air semua pemandunya cerita, kalau banyak hape yg nyemplung dan ga bs diambil lg karena dalem banget. Ada spot dimana bisa lompat juga dr sejenis tebing. Tapi waktu itu ga nyobain karena takut, temen2 yg lain nyobain. Seru pokoknya.	We went here during Ramadan, so it was a bit quieter and there were no queues. The guides were very experienced, friendly, and loved to tell stories, so they knew the story. If you want to bring your phone, make sure it's in a pocket and can be worn around your neck, because the guides were telling stories throughout the trip. If you had too many phones, they'd fall in and you couldn't retrieve them because it was so deep. There's also a spot where you can jump off a cliff. But I didn't try it because I was afraid, but my friends did. It was really fun.
2	Mohon penertiban calo dan markup tarif, khusus calo ditolak halus tp tetap mengarahkan	Please regulate brokers and mark up rates, especially brokers, politely reject them but still direct them
3	tingkatkan terus @guapindul_jogja 🍊 seruuu 🍊	keep improving @guapindul_jogja 🍊 so much fun 🍊

4.2.2. Filtering

Tahap *preprocessing* yang kedua adalah *filtering*. Tahap ini bertujuan untuk menghilangkan kata atau karakter yang tidak memiliki relevansi terhadap pemodelan topik atau analisis sentimen. Kata atau karakter yang disaring pada tahap ini meliputi penghapusan *emoticon*, *mention username*, *hashtag*, dan karakter *non-alphabet*. Keseluruhan tahap *filtering* ini dilakukan dengan *function* dari *Microsoft Excel*. Berikut merupakan tabel dari *function Microsoft Excel* yang digunakan:

Tabel 4. 2 *Function filtering Excel*

No	Tujuan	Function Excel
1	Penghapusan <i>Emoticon</i>	= <i>RemoveEmoji</i> (G2)
2	Penghapusan <i>mention</i> <i>username</i> dan <i>hashtag</i>	Penghapusan dilakukan secara manual dengan memanfaatkan fitur “ <i>sort by</i> ” dan “ <i>Conditional Formatting</i> ”
3	Karakter non- <i>alphabet</i>	=LET(teks; I2; daftar; {"~";"1";"2";"3";"4";"5";"6";"7";"8";"9";"0";"-";"=";"[";"]";"\";";";"'" ;";"." ;"/";"!" ;"@";"#";"\$";"%";"^";"&";"*";"(" ;")";"_" ;"+" ;"{" ;"}";" " ;":" ;"<" ;">" ;"?" } ; REDUCE (teks; daftar; LAMBDA (a; b; SUBSTITUTE(a;b;""))))

Berdasarkan hasil *function* tersebut, didapatkan *text* yang sudah dilakukan tahap *filtering* sebagai berikut:

Tabel 4. 3 Hasil *filtering*

No	Text Translate	Filtering
1	We went here during Ramadan, so it was a bit quieter and there were no queues. The guides were very experienced, friendly, and loved to tell stories, so they knew the story. If you want to bring your phone, make sure it's in a pocket and can be worn around your neck, because the guides were telling stories throughout the trip. If you had too many phones, they'd fall in and you couldn't retrieve them because it was so deep. There's also a spot where you can jump off a cliff. But I didn't try it because I was afraid, but my friends did. It was really fun.	We went here during Ramadan so it was a bit quieter and there were no queues The guides were very experienced friendly and loved to tell stories so they knew the story If you want to bring your phone make sure its in a pocket and can be worn around your neck because the guides were telling stories throughout the trip If you had too many phones theyd fall in and you couldnt retrieve them because it was so deep Theres also a spot where you can jump off a cliff But I didnt try it because I was afraid but my friends did It was really fun

No	<i>Text Translate</i>	<i>Filtering</i>
2	Please regulate brokers and mark up rates, especially brokers, politely reject them but still direct them	Please regulate brokers and mark up rates especially brokers politely reject them but still direct them
5	keep improving @guapindul_jogja 🍌 so much fun 🍌	keep improving so much fun

4.2.3. Case Folding

Tahap *preprocessing data* selanjutnya adalah melakukan *case folding*, yaitu mengubah semua karakter atau huruf kapital menjadi huruf kecil pada teks yang sudah dilakukan *filtering*. Tujuan dari tahap ini adalah untuk menghindari nilai bias dari hasil pemodelan topik atau analisis sentimen yang akan dilakukan. Tahapan *casefolding* ini dilakukan dengan menggunakan *function Microsoft Excel*, yaitu “=Lower(J2)”. Berikut ini merupakan hasil dari *preprocessing data* tahap *case folding* yang telah dilakukan:

Tabel 4. 4 Hasil *case folding*

No	<i>Filtering</i>	<i>Case Folding</i>
1	We went here during Ramadan so it was a bit quieter and there were no queues The guides were very experienced friendly and loved to tell stories so they knew the story If you want to bring your phone make sure its in a pocket and can be worn around your neck because the guides were telling stories throughout the trip If you had too many phones theyd fall in and you couldnt retrieve them because it was so deep Theres also a spot where you can jump off a cliff But I didnt try it because I was afraid but my friends did It was really fun	we went here during ramadan so it was a bit quieter and there were no queues the guides were very experienced friendly and loved to tell stories so they knew the story if you want to bring your phone make sure its in a pocket and can be worn around your neck because the guides were telling stories throughout the trip if you had too many phones theyd fall in and you couldnt retrieve them because it was so deep theres also a spot where you can jump off a cliff but i didnt try it because i was afraid but my friends did it was really fun

No	<i>Filtering</i>	<i>Case Folding</i>
2	Please regulate brokers and mark up rates especially brokers politely reject them but still direct them	please regulate brokers and mark up rates especially brokers politely reject them but still direct them
3	keep improving so much fun	keep improving so much fun

4.2.4. Stopwords Removal

Tahap *stopwords removal* bertujuan untuk menghapus kata-kata yang tidak memiliki makna penting dalam analisis (*stopwords*), seperti kata hubung dan kata umum lainnya. Pada tahap ini, kembali menggunakan *function* yang tersedia di *Microsoft Excel*. Berikut merupakan *function* yang digunakan pada penelitian ini:

Tabel 4. 5 *Function Excel* untuk *stopwords removal*

Tujuan	<i>Function Excel</i>
<i>Stopword Removal</i>	=TEXTJOIN(" "; TRUE; FILTER (TEXTSPLIT (L2; " "); ISNA (MATCH (TEXTSPLIT (L2; " "); {"a"; "an"; "the"; "in"; "at"; "on"; "of"; "off"; "if"; "for"; "with"; "by"; "to"; "didn't"; "as"; "doesn't"; "aren't"; "wasn't"; "werent"; "isn't"; "and"; "but"; "or"; "so"; "yet"; "is"; "am"; "are"; "was"; "were"; "be"; "been"; "being"; "have"; "has"; "had"; "do"; "does"; "did"; "you"; "your"; "there"; "they"; "just"; "also"; "dont"; "will"; "can"; "get"; "still"; "much"; "many"; "very"; "from"; "its"; "i"; "get"; "that"; "this"; "here"; "no"; "not"; "any"; "im"; "we"; "their"; "yours"; "them"; "us"; "our"; "who"; "what"; "why"; "it"; "how"; "min"; "ok"; "thats"; "too"; "wow"; "k"; "rp"; "he"; "she"; "her"; "him"; "my"; "very"; "still"; "u"; "ur"; "pls"; "plz"; "haha"; "hehe"; "wkwk"; "lol"; "cuz"; "cuzzz"; "go"; "come"}; 0))))

Berdasarkan *function Microsoft Excel* tersebut, maka didapatkan perubahan dari data *case folding* menjadi data *stopword removal* sebagai berikut:

Tabel 4. 6 Hasil *stopword removal*

No	Case Folding	Stopword Removal
1	we went here during ramadan so it was a bit quieter and there were no queues the guides were very experienced friendly and loved to tell stories so they knew the story if you want to bring your phone make sure its in a pocket and can be worn around your neck because the guides were telling stories throughout the trip if you had too many phones theyd fall in and you couldnt retrieve them because it was so deep theres also a spot where you can jump off a cliff but i didnt try it because i was afraid but my friends did it was really fun	went during ramadan bit quieter queues guides experienced friendly loved tell stories knew story want bring phone make sure pocket worn around neck because guides telling stories throughout trip phones theyd fall couldnt retrieve because deep theres spot where jump cliff try because afraid friends really fun
2	please regulate brokers and mark up rates especially brokers politely reject them but still direct them	please regulate brokers mark up rates especially brokers politely reject direct
3	keep improving so much fun	keep improving fun

4.2.5. Tokenisasi

Tahap tokenisasi merupakan proses pemecahan teks menjadi unit-unit kata (token) sehingga dapat dianalisis lebih lanjut dalam proses *text mining*. Pada tahap ini, teks hasil *stopwords removal* diubah menjadi kumpulan kata individual dengan memisahkan setiap kata dalam kalimat.

Proses tokenisasi dalam penelitian ini dilakukan menggunakan bahasa pemrograman *Python* dengan bantuan *library* NLTK, khususnya fungsi *word_tokenize*. Fungsi ini digunakan untuk mengubah teks menjadi *list* kata sehingga setiap kata dapat diproses sebagai unit analisis pada tahap selanjutnya. Berikut merupakan contoh perubahan data dari *stopword removal* menjadi data yang ditokenisasikan:

Tabel 4. 7 Hasil tokenisasi

No	Stopword Removal	Tokenisasi
1	went during ramadan bit quieter queues guides experienced friendly loved tell stories knew story want bring phone make sure pocket worn around neck because guides telling stories throughout trip phones theyd fall couldnt retrieve because deep theres spot where jump cliff try because afraid friends really fun	['went', 'during', 'ramadan', 'bit', 'quieter', 'queues', 'guides', 'experienced', 'friendly', 'loved', 'tell', 'stories', 'knew', 'story', 'want', 'bring', 'phone', 'make', 'sure', 'pocket', 'worn', 'around', 'neck', 'because', 'guides', 'telling', 'stories', 'throughout', 'trip', 'phones', 'theyd', 'fall', 'couldnt', 'retrieve', 'because', 'deep', 'theres', 'spot', 'where', 'jump', 'cliff', 'try', 'because', 'afraid', 'friends', 'really', 'fun']
2	please regulate brokers mark up rates especially brokers politely reject direct	['please', 'regulate', 'brokers', 'mark', 'up', 'rates', 'especially', 'brokers', 'politely', 'reject', 'direct']
3	keep improving fun	['keep', 'improving', 'fun']

4.2.6. Stemming

Tahap stemming merupakan proses mengubah kata hasil tokenisasi menjadi bentuk dasar (*root word*) untuk mengurangi variasi kata yang memiliki makna serupa. Proses ini bertujuan untuk meningkatkan kualitas analisis dengan menyederhanakan representasi kata dalam dataset.

Pada penelitian ini, proses stemming dilakukan menggunakan library NLTK dengan metode *PorterStemmer* (*from nltk.stem import PorterStemmer*). Metode ini bekerja dengan menghapus imbuhan pada kata sehingga diperoleh bentuk dasar yang lebih konsisten dan siap digunakan dalam proses pemodelan topik menggunakan LDA. Berikut ini merupakan contoh perubahan data dari tokenisasi menjadi data yang telah dilakukan *stemming*:

Tabel 4. 8 Hasil *stemming*

No	Tokenisasi	Stemming
1	['went', 'during', 'ramadan', 'bit', 'quieter', 'queues', 'guides', 'experienced', 'friendly', 'loved', 'tell', 'stories', 'knew', 'story', 'want', 'bring', 'phone', 'make', 'sure',	['went', 'dure', 'ramadan', 'bit', 'quieter', 'queue', 'guid', 'experienc', 'friendli', 'love', 'tell', 'stori', 'knew', 'stori', 'want', 'bring', 'phone', 'make', 'sure', 'pocket', 'worn',

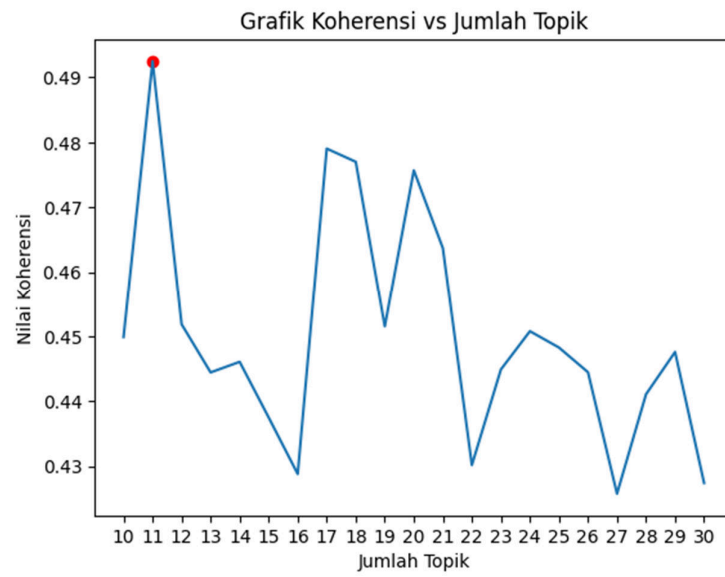
No	Tokenisasi	Stemming
	'pocket', 'worn', 'around', 'neck', 'because', 'guides', 'telling', 'stories', 'throughout', 'trip', 'phones', 'theyd', 'fall', 'couldnt', 'retrieve', 'because', 'deep', 'theres', 'spot', 'where', 'jump', 'cliff', 'try', 'because', 'afraid', 'friends', 'really', 'fun']	'around', 'neck', 'becaus', 'guid', 'tell', 'stori', 'throughout', 'trip', 'phone', 'theyd', 'fall', 'couldnt', 'retriev', 'becaus', 'deep', 'there', 'spot', 'where', 'jump', 'cliff', 'tri', 'becaus', 'afraid', 'friend', 'realli', 'fun']
2	['please', 'regulate', 'brokers', 'mark', 'up', 'rates', 'especially', 'brokers', 'politely', 'reject', 'direct']	['pleas', 'regul', 'broker', 'mark', 'up', 'rate', 'especi', 'broker', 'polit', 'reject', 'direct']
3	['keep', 'improving', 'fun']	['keep', 'improv', 'fun']

4.3 Latent Dirichlet Allocation (LDA)

Pada tahap ini dilakukan pemodelan topik menggunakan metode *Latent Dirichlet Allocation* (LDA) untuk mengidentifikasi tema-tema utama yang muncul dalam ulasan pengunjung wisata Goa Pindul. Proses pemodelan dilakukan menggunakan bahasa pemrograman *Python* dengan memanfaatkan *library gensim*, yang memungkinkan pembentukan model topik secara probabilistik berdasarkan distribusi kata dalam dokumen.

4.3.1. Pemodelan Topik Data Google Maps

Pada proses ini dilakukan pengujian beberapa variasi jumlah topik (k) untuk menentukan jumlah topik yang paling optimal untuk ulasan *Google Maps*. Penentuan jumlah topik dilakukan dengan menghitung nilai *coherence score* pada setiap variasi jumlah topik, kemudian memilih nilai dengan skor koherensi tertinggi sebagai jumlah topik terbaik. Berdasarkan hasil pengujian, diperoleh jumlah topik optimal sebanyak 11 topik dengan nilai koherensi sebesar 0,492.



Gambar 4. 1 Grafik koherensi LDA ulasan *Google Maps*

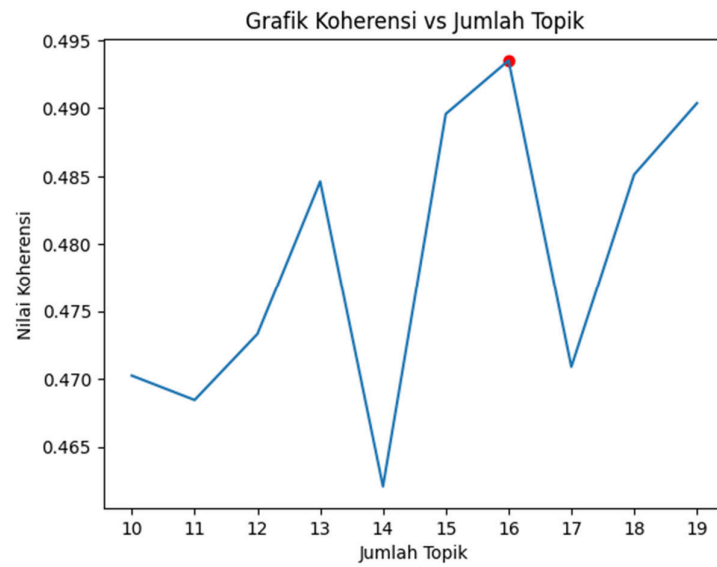
Berikut merupakan hasil pemodelan topik beserta label interpretasi untuk masing-masing topik dari data ulasan *Google Maps*:

Tabel 4. 9 *Label topik ulasan Google Maps*

Topik	<i>Label Topik</i>
0	Keseruan Wisata Goa Pindul
1	Aktivitas <i>Rafting</i>
2	Kebersihan Fasilitas dan Keindahan Alam
3	Kepuasan dan Pengalaman Wisata
4	Eksplorasi Wisata <i>Cave Tubing</i>
5	Pengalaman Petualangan Wisata
6	Keseruan <i>Tour</i> dan Keramahan Pemandu
7	Liburan di Goa Pindul
8	Pengalaman Pertama ke Goa Pindul
9	Pelayanan Wisata
10	Kepuasan Layanan untuk Keluarga

4.3.2. *Pemodelan Topik Data Instagram*

Pada proses ini dilakukan pengujian beberapa variasi jumlah topik (k) untuk menentukan jumlah topik yang paling optimal untuk ulasan *Instagram*. Penentuan jumlah topik dilakukan dengan menghitung nilai *coherence score* pada setiap variasi jumlah topik, kemudian memilih nilai dengan skor koherensi tertinggi sebagai jumlah topik terbaik. Berdasarkan hasil pengujian, diperoleh jumlah topik optimal sebanyak 16 topik dengan nilai koherensi sebesar 0,493.



Gambar 4. 2 Grafik koherensi LDA ulasan *Instagram*

Berikut merupakan hasil pemodelan topik beserta label interpretasi untuk masing-masing topik dari data ulasan *Instagram*:

Tabel 4. 10 *Label* topik ulasan *Instagram*

Topik	Label
0	Keamanan Anak dan Batas Usia
1	Aktivitas Seru dan Rafting
2	Tiket dan Akses Masuk
3	Aktivitas Wisata Keluarga
4	Kekhawatiran dan Risiko Wisata
5	Harapan Wisatawan Terhadap Goa Pindul
6	Keindahan Sungai dan Wisata Air
7	Keluhan dan Mitigasi Wisata
8	Ucapan Terima Kasih Pengunjung
9	Komentar Santai Pengunjung
10	Pengalaman Keluarga dan Hiburan
11	Kepuasan Wisatawan Goa Pindul
12	Harapan dan Aktivitas Wisata
13	Interaksi dan Pelayanan Admin
14	Informasi Tiket dan Waktu
15	Reservasi dan Administrasi Wisata

4.4 *Valence Aware Dictionary and sEntiment Researnor (VADER)*

4.4.1. *Analisis Sentimen Data Google Maps*

Analisis sentimen pada data ulasan *Google Maps* dilakukan menggunakan metode berbasis leksikon, yaitu *VADER SentimentIntensityAnalyzer* yang tersedia dalam *library* NLTK. Metode ini digunakan untuk menghitung skor sentimen pada setiap ulasan dalam bentuk nilai positif, negatif, netral, serta skor gabungan (*compound score*). Dalam penelitian ini, *compound score* digunakan sebagai nilai utama karena mampu merepresentasikan kecenderungan sentimen secara keseluruhan dalam satu nilai numerik.

Setelah diperoleh nilai *compound score*, dilakukan proses klasifikasi sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral. Klasifikasi dilakukan dengan menggunakan batasan nilai, yaitu skor $\geq 0,05$ dikategorikan sebagai sentimen positif, skor $\leq -0,05$ sebagai sentimen negatif, dan nilai di antara kedua batas tersebut dikategorikan sebagai netral. Hasil analisis ini

digunakan untuk menggambarkan kecenderungan opini pengunjung terhadap wisata Goa Pindul berdasarkan data ulasan *Google Maps*.

Tabel 4. 11 *Sentiment score* ulasan *Google Maps*

Teks	Text Translate	sentiment_score
0	It's really good	0.4927
1	no one is selling cigarettes	-0.2960
2	Best for those who want to play in the water a...	0.9612
3	Great nature. Exciting experience. Deceng tour...	-0.0258
4	Recommendations for families for year-end holi...	0.3818

4.4.2. Analisis Sentimen Data Instagram

Analisis sentimen pada data komentar Instagram dilakukan menggunakan metode yang sama, yaitu *VADER SentimentIntensityAnalyzer*. Setiap komentar dianalisis untuk memperoleh nilai *compound score* yang mencerminkan kecenderungan opini pengguna terhadap wisata Goa Pindul.

Selanjutnya, nilai *compound score* tersebut diklasifikasikan ke dalam kategori positif, negatif, dan netral menggunakan ambang batas yang sama, yaitu $\geq 0,05$ untuk positif, $\leq -0,05$ untuk negatif, dan di antaranya sebagai netral. Hasil klasifikasi ini digunakan untuk menggambarkan kecenderungan sentimen pengunjung pada *platform* Instagram berdasarkan komentar yang diperoleh.

Tabel 4. 12 *Sentimen score* ulasan *Instagram*

Teks	Text Translate	sentiment_score
0	Bonus bonus	0.7906
1	Cool	0.3182
2	Min, can a 5 year old child join?	0.296
3	I'm here, Min. I really miss Jogja and want to...	-0.1513
4	It was so much fun and the atmosphere was so s...	0.7996

4.5 Importance-Performance Analysis (IPA) Matrix

Pada tahap ini dilakukan analisis menggunakan metode *Importance Performance Analysis* (IPA) untuk menentukan prioritas perbaikan berdasarkan persepsi pengunjung terhadap wisata Goa Pindul. Metode IPA digunakan dengan memetakan atribut pelayanan ke dalam dua dimensi, yaitu tingkat kepentingan (*importance*) dan tingkat kinerja (*performance*), yang kemudian divisualisasikan dalam bentuk diagram kartesius.

4.5.1. Penentuan Nilai Importance

Nilai *importance* dalam penelitian ini diperoleh dari nilai bobot kepentingan (*importance weight*) pada setiap topik yang dihasilkan dari pemodelan *Latent Dirichlet Allocation* (LDA). Langkah awal dalam menentukan nilai *importance* ini diperlukan total bobot kepentingan untuk masing-masing topik,

$$WI_t = \sum_{i=0}^n Weight_i \quad (4.1)$$

$$Importance\ score = \frac{WI_t - WI_{t_{min}}}{WI_{t_{max}} - WI_{t_{min}}} \quad (4.2)$$

Berikut merupakan hasil perhitungan total bobot kepentingan untuk topik pada ulasan *Google Maps* dan *Instagram*:

Tabel 4. 13 Total bobot kepentingan topik ulasan

Topik	Sumber	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	WI _t
0	GMAPS	0.121	0.06	0.041	0.041	0.04	0.034	0.017	0.015	0.014	0.013	0.396
1	GMAPS	0.09	0.042	0.03	0.02	0.015	0.013	0.012	0.012	0.011	0.01	0.255
2	GMAPS	0.121	0.04	0.03	0.029	0.022	0.017	0.017	0.014	0.014	0.014	0.318
3	GMAPS	0.048	0.022	0.02	0.018	0.017	0.016	0.015	0.015	0.013	0.013	0.197
4	GMAPS	0.09	0.028	0.023	0.023	0.023	0.017	0.014	0.011	0.011	0.01	0.25
5	GMAPS	0.039	0.036	0.024	0.024	0.023	0.021	0.019	0.019	0.019	0.018	0.242
6	GMAPS	0.085	0.074	0.054	0.053	0.047	0.045	0.034	0.024	0.023	0.019	0.458
7	GMAPS	0.061	0.045	0.037	0.031	0.025	0.022	0.019	0.017	0.014	0.014	0.285
8	GMAPS	0.027	0.021	0.013	0.013	0.012	0.012	0.011	0.011	0.011	0.01	0.141
9	GMAPS	0.051	0.046	0.019	0.017	0.014	0.014	0.014	0.013	0.012	0.012	0.212
10	GMAPS	0.106	0.069	0.039	0.031	0.031	0.026	0.025	0.022	0.021	0.018	0.388
0	IG	0.059	0.056	0.045	0.037	0.035	0.035	0.032	0.032	0.027	0.022	0.38
1	IG	0.033	0.032	0.031	0.022	0.02	0.019	0.019	0.017	0.017	0.015	0.225
2	IG	0.055	0.027	0.027	0.027	0.02	0.02	0.016	0.016	0.012	0.012	0.232
3	IG	0.118	0.04	0.03	0.028	0.023	0.016	0.016	0.013	0.013	0.012	0.309
4	IG	0.037	0.035	0.029	0.028	0.022	0.016	0.013	0.013	0.013	0.013	0.219
5	IG	0.083	0.082	0.025	0.014	0.013	0.013	0.011	0.011	0.009	0.009	0.27

Topik	Sumber	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	WI _t
6	IG	0.035	0.026	0.018	0.017	0.016	0.016	0.015	0.015	0.012	0.012	0.182
7	IG	0.05	0.041	0.017	0.012	0.012	0.012	0.012	0.011	0.011	0.009	0.187
8	IG	0.057	0.032	0.02	0.02	0.014	0.01	0.01	0.01	0.01	0.01	0.193
9	IG	0.022	0.02	0.016	0.014	0.013	0.013	0.013	0.009	0.009	0.009	0.138
10	IG	0.111	0.036	0.019	0.016	0.015	0.014	0.014	0.012	0.011	0.01	0.258
11	IG	0.04	0.038	0.034	0.031	0.019	0.017	0.016	0.016	0.016	0.016	0.243
12	IG	0.029	0.024	0.02	0.02	0.02	0.018	0.015	0.015	0.015	0.015	0.191
13	IG	0.061	0.043	0.038	0.033	0.021	0.019	0.019	0.018	0.018	0.013	0.283
14	IG	0.021	0.017	0.012	0.012	0.012	0.011	0.01	0.009	0.009	0.009	0.122
15	IG	0.022	0.019	0.018	0.018	0.016	0.016	0.016	0.013	0.011	0.011	0.16

Dari hasil total bobot kepentingan tiap-tiap topik, maka normalisasi *importance score* dapat ditentukan dengan rumus:

$$Importance\ score = \frac{WI_t - WI_{t_{min}}}{WI_{t_{max}} - WI_{t_{min}}} \quad (4.2)$$

Berikut merupakan *importance score* dari tiap-tiap topik ulasan *Google Maps* berdasarkan nilai WI_t yang didapatkan sebelumnya:

Tabel 4. 14 *Importance score* topik ulasan *Google Maps*

Topik	Sumber	WI _t	Importance Score
0	GMAPS	0.396	0.804416
1	GMAPS	0.255	0.359621
2	GMAPS	0.318	0.55836
3	GMAPS	0.197	0.176656
4	GMAPS	0.25	0.343849
5	GMAPS	0.242	0.318612
6	GMAPS	0.458	1
7	GMAPS	0.285	0.454259
8	GMAPS	0.141	0
9	GMAPS	0.212	0.223975
10	GMAPS	0.388	0.77918
0	IG	0.218	0.0845
1	IG	0.194	0.2575
2	IG	0.116	0.4636
3	IG	0.326	0.0000

Topik	Sumber	WI_t	Importance Score
4	IG	0.185	0.5943
5	IG	0.138	0.6606
6	IG	0.181	0.8320
7	IG	0.206	0.8340
8	IG	0.284	0.1648
9	IG	0.246	1.0000
10	IG	0.286	0.0709
11	IG	0.279	0.0845
12	IG	0.199	0.2575
13	IG	0.217	0.4636
14	IG	0.205	0.0000
15	IG	0.127	0.5943

4.5.2. Penentuan Nilai Performance

Nilai *performance* diperoleh dari rata-rata compound score hasil analisis sentimen menggunakan VADER *SentimentIntensityAnalyzer* pada setiap topik. Untuk mendapatkan nilai *performance* berdasarkan hasil nilai *compound score* berlabel negatif yang telah didapatkan adalah dengan rumus,

$$WS_t = \frac{\sum S_i}{n_k} \quad (4.3)$$

$$performance\ score = \frac{WS_t - WS_{t_{min}}}{WS_{t_{max}} - WS_{t_{min}}} \quad (4.4)$$

Berikut merupakan hasil *performance score* untuk tiap-tiap topik yang didapatkan:

Tabel 4. 15 *Performance score IPA matrix*

Topik	Sumber	WSt	Performance Score
0	GMAPS	-0.3857	0.5313
1	GMAPS	-0.4729	0.0948

Topik	Sumber	WSt	Performance Score
2	GMAPS	-0.4633	0.1425
3	GMAPS	-0.3211	0.8551
4	GMAPS	-0.3464	0.7282
5	GMAPS	-0.3581	0.6697
6	GMAPS	-0.3138	0.8916
7	GMAPS	-0.2922	1.0000
8	GMAPS	-0.4367	0.2758
9	GMAPS	-0.4918	1.0000
10	GMAPS	-0.3514	0.7035
0	IG	-0.2732	0.6319
1	IG	-0.3649	0.4653
2	IG	-0.5299	0.1656
3	IG	-0.0706	1.0000
4	IG	-0.2671	0.6429
5	IG	-0.2600	0.6559
6	IG	-0.2868	0.6072
7	IG	-0.5092	0.2031
8	IG	-0.3831	0.4324
9	IG	-0.3375	0.5150
10	IG	-0.6210	0.0000
11	IG	-0.5384	0.1502
12	IG	-0.5076	0.2060
13	IG	-0.3626	0.4695
14	IG	-0.3600	0.4742
15	IG	-0.2732	0.6319

4.5.3. Visualisasi Importance-Performance Analysis (IPA) Matrix

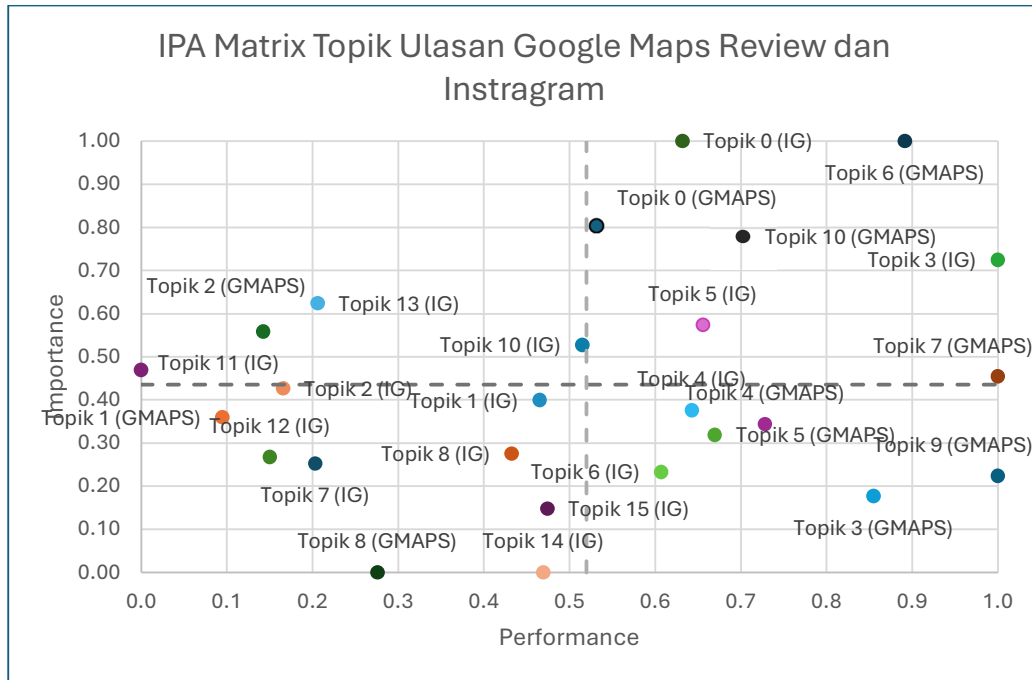
Nilai *performance* dan *importance* yang telah didapatkan, selanjutnya dilakukan perhitungan nilai rata-rata keseluruhan untuk *importance* dan *performance* yang berfungsi sebagai titik potong (*grand mean*) pada diagram kartesius. Berdasarkan hasil perhitungan, diperoleh nilai rata-rata *importance* sebesar 0.4213 dan nilai rata-rata *performance* sebesar 0.4575. Kedua nilai tersebut digunakan sebagai garis pembagi yang membentuk empat kuadran dalam diagram IPA.

Tabel 4. 16 *Performance dan importance score*

Topik	Sumber	Performance (x-axis)	Importance (y-axis)
0	GMAPS	0.5313	0.8044
1	GMAPS	0.0948	0.3596
2	GMAPS	0.1425	0.5584
3	GMAPS	0.8551	0.1767
4	GMAPS	0.7282	0.3438

Topik	Sumber	Performance (x-axis)	Importance (y-axis)
5	GMAPS	0.6697	0.3186
6	GMAPS	0.8916	1.0000
7	GMAPS	1.0000	0.4543
8	GMAPS	0.2758	0.0000
9	GMAPS	1.0000	0.2240
10	GMAPS	0.7035	0.7792
0	IG	0.6319	1.0000
1	IG	0.4653	0.3992
2	IG	0.1656	0.4264
3	IG	1.0000	0.7248
4	IG	0.6429	0.3760
5	IG	0.6559	0.5736
6	IG	0.6072	0.2326
7	IG	0.2031	0.2519
8	IG	0.4324	0.2752
9	IG	0.5150	0.5271
10	IG	0.0000	0.4690
11	IG	0.1502	0.2674
12	IG	0.2060	0.6240
13	IG	0.4695	0.0000
14	IG	0.4742	0.1473
15	IG	0.5313	0.8044
Rata-rata		0.4351	0.5197

Setiap topik kemudian dipetakan ke dalam diagram kartesius sesuai dengan koordinat nilai *importance* dan *performance* yang dimiliki. Hasil pemetaan ini menghasilkan posisi masing-masing topik dalam empat kuadran IPA, yaitu Kuadran I (*Concentrate Here*), Kuadran II (*Keep Up the Good Work*), Kuadran III (*Low Priority*), dan Kuadran IV (*Possible Overkill*). Visualisasi ini bertujuan untuk memberikan gambaran distribusi atribut berdasarkan tingkat kepentingan dan kinerja, yang selanjutnya akan digunakan sebagai dasar dalam tahap analisis dan pembahasan. Berikut merupakan diagram kartesius hasil IPA dari ulasan *Google Maps* dan *Instagram*:



Gambar 4. 3 IPA *matrix* ulasan Goa Pindul

BAB V

HASIL DAN PEMBAHASAN

5.1. Hasil *Latent Dirichlet Allocation* (LDA)

Pemodelan topik memiliki tujuan untuk mengelompokkan ulasan teks, baik pada *Google Maps* maupun *Instagram* dari pengunjung Goa Pindul ke dalam beberapa topik utama. Metode pemodelan topik pada penelitian ini menggunakan metode *Latent Dirichlet Allocation* (LDA) yang merupakan suatu metode yang dapat melakukan pengelompokkan data kedalam beberapa topik dalam satu dokumennya.

Pada ulasan dari *Google Maps*, telah dilakukan pengelompokkan topik pada tahap pengolahan data sebelumnya. Berdasarkan tahap tersebut, didapatkan nilai koherensi sebesar 0,492 pada total 11 topik. Hal ini menunjukkan bahwa terdapat 11 topik yang menjadi pembahasan utama pada ulasan *Google Maps* yang dilakukan oleh pengunjung Goa Pindul selama periode Januari 2020 – Februari 2026. Topik-topik utama ini dapat merepresentasikan tingkat kepentingan yang dapat menjadi pertimbangan bagi pengelola Goa Pindul. Berikut merupakan interpretasi dari tiap-tiap topik yang terbentuk pada ulasan *Google Maps* Goa Pindul:

Tabel 5. 1 Interpretasi topik pada ulasan *Google Maps*

Topik	Keywords Dominan	Label Topik	Interpretasi
0	<i>Excit, cave, pindul, fun, tourism, water, explor, best, kidul, natur</i>	Keseruan Wisata Goa Pindul	Topik ini menggambarkan pengalaman wisata <i>cave tubing</i> yang memberikan sensasi menyenangkan bagi pengunjung melalui aktivitas menyusuri goa dan aliran sungai.
1	<i>Cave, pindul, river, water, raft, stalactit, beauti, stalagmit, explor, flow</i>	Aktivitas <i>Rafting</i>	Topik ini merepresentasikan aktivitas eksplorasi Goa Pindul yang melibatkan elemen alam seperti sungai, stalaktit, dan stalagmit sebagai daya tarik utama.

Topik	Keywords Dominan	Label Topik	Interpretasi
2	<i>Good, natur, tourism, tourist, place, locat, servic, kidul, interest, area</i>	Keindahan Alam	Topik ini menunjukkan persepsi pengunjung terhadap lokasi wisata, termasuk aspek lingkungan, fasilitas, dan layanan yang diberikan.
3	<i>Awesom, good, bad, would, more, tire, cave, satisfactori, long, ive</i>	Kepuasan dan Pengalaman Wisata	Topik ini mencerminkan penilaian kepuasan pengunjung yang bersifat umum, baik berupa pengalaman positif maupun negatif selama berkunjung.
4	<i>Cave, pindul, tube, explor, tourist, attract, take, locat, enter, quit</i>	Eksplorasi Wisata <i>Cave Tubing</i>	Topik ini berkaitan dengan aktivitas utama wisata Goa Pindul, seperti <i>cave tubing</i> dan interaksi langsung dengan lingkungan gua.
5	<i>Amaz, like, guid, place, experi, adventur, fun, those, cave, pleasant</i>	Pengalaman Petualangan Wisata	Topik ini menggambarkan pengalaman petualangan yang dirasakan pengunjung selama mengikuti aktivitas wisata.
6	<i>Place, fun, cool, nice, guid, realli, friendli, tour, water, comfort</i>	Keseruan <i>Tour</i> dan Keramahan Pemandu	Topik ini menunjukkan aspek kenyamanan dan pelayanan yang dirasakan pengunjung selama menikmati aktivitas wisata.
7	<i>Beauti, water, holiday, dure, season, natur, cave, experi, fun, crowd</i>	Liburan di Goa Pindul	Topik ini menggambarkan keindahan alam yang menjadi daya tarik wisata, seperti kondisi lingkungan dan suasana alam sekitar.

Topik	Keywords Dominan	Label Topik	Interpretasi
8	<i>Cave, because, fun, when, guid, lot, peopl, water, first, bring</i>	Pengalaman Pertama yang Menyenangkan	Topik ini berkaitan dengan cerita pengalaman pengunjung yang pertama kali ke Goa Pindul
9	<i>Cave, pindul, guid, river, goa, tube, idr, friendli, tour, servic</i>	Pelayanan Wisata	Topik ini menunjukkan aspek layanan teknis pelayanan wisata.
10	<i>Great, servic, adventur, famili, satisfi, awesom, recreat, happi, travel, top</i>	Kepuasan Layanan untuk Keluarga	Topik ini mencerminkan tingkat kepuasan pengunjung terhadap layanan dan pengalaman rekreasi secara keseluruhan.

Selanjutnya, ulasan *Instagram* juga dilakukan pemodelan topik dengan metode *Latent Dirichlet Allocation* (LDA). Tujuan dari pemodelan topik pada ulasan *Instagram* untuk mengetahui topik-topik utama yang menjadi pembahasan pengunjung tempat wisata Goa Pindul dalam periode Januari 2020 – Februari 2026. Hasil dari pemodelan topik dengan metode LDA pada ulasan *Instagram* adalah terdiri dari 16 topik yang terbentuk. Nilai koherensi dari ulasan *Instagram* ini adalah 0,493. Berikut merupakan tabel untuk hasil interpretasi dari tiap-tiap topik yang terbentuk pada ulasan *Instagram* di akun @goapindul.id dan @guapindul_jogja:

Tabel 5. 2 Interpretasi topik pada ulasan Instagram

Topik	Keywords Dominan	Label Topik	Interpretasi
0	<i>Children, year, old, safe, age, child, join, awesom, si, minimum</i>	Keamanan Anak dan Batas Usia	Topik ini membahas perhatian wisatawan terhadap aspek keamanan anak dalam aktivitas wisata Goa Pindul. Kata seperti children, safe, age, dan minimum menunjukkan adanya diskusi mengenai batas usia minimum, keamanan anak-anak, serta kelayakan wisata

Topik	Keywords Dominan	Label Topik	Interpretasi
			keluarga untuk berbagai kelompok umur.
1	<i>Tri, fun, want, excit, realli, me, raft, right, ive, befor</i>	Aktivitas Seru dan Rafting	Topik ini merepresentasikan aktivitas wisata yang bersifat menyenangkan dan menantang, seperti kegiatan rafting atau permainan air.
2	<i>Want, now, peopl, ticket, entranc, close, play, si, money, th</i>	Tiket dan Akses Masuk	Topik ini berkaitan dengan informasi mengenai tiket masuk, akses wisata, serta biaya yang perlu diketahui oleh pengunjung.
3	<i>Fun, want, kid, realli, jogja, oh, god, again, cave, join</i>	Aktivitas Wisata Keluarga	Topik ini menunjukkan aktivitas wisata yang melibatkan keluarga, khususnya kegiatan yang dapat dinikmati bersama anak-anak.
4	<i>Want, miss, realli, again, bring, possibl, ive, jogja, safe, scari</i>	Kekhawatiran dan Risiko Wisata	Topik ini mencerminkan kekhawatiran atau risiko yang dirasakan pengunjung selama melakukan aktivitas wisata.
5	<i>Cave, pindul, tourist, goa, hope, tour, guid, want, best, trip</i>	Harapan Wisatawan Terhadap Goa Pindul	Topik ini menggambarkan harapan dan ekspektasi wisatawan terhadap pengalaman wisata di Goa Pindul. Kata hope, best, trip, dan guid menunjukkan adanya keinginan wisatawan untuk memperoleh pengalaman wisata yang memuaskan dengan pelayanan yang baik.

Topik	Keywords Dominan	Label Topik	Interpretasi
6	<i>Water, beauti, time, river, long, when, prevent, bro, because, fun</i>	Keindahan Sungai dan Wisata Air	Topik ini membahas daya tarik alam berupa sungai dan wisata air di Goa Pindul. Kata <i>water, beauti, river, dan fun</i> menunjukkan bahwa wisatawan menikmati keindahan alam sungai serta suasana wisata air yang menyenangkan.
7	<i>Cave, pindul, pleas, better, disast, mitig, video, hope, locat, jogja</i>	Keluhan dan Mitigasi Wisata	Topik ini berkaitan dengan keluhan atau masalah yang dialami pengunjung, serta upaya mitigasi atau perbaikan yang dilakukan.
8	<i>Thank, brother, more, your, si, like, raini, season, went, welcom</i>	Ucapan Terima Kasih Pengunjung	Topik ini mencerminkan ungkapan apresiasi atau ucapan terima kasih dari pengunjung setelah melakukan aktivitas wisata.
9	<i>Let, know, like, si, thousand, theyr, hahahaha, natur, life, travel</i>	Komentar Santai Pengunjung	Topik ini menunjukkan komentar santai atau informal dari pengunjung yang tidak secara langsung berkaitan dengan aspek utama wisata.
10	<i>Cool, realli, si, want, famili, nice, fun, bro, all, like</i>	Pengalaman Keluarga dan Hiburan	Topik ini menggambarkan pengalaman wisata yang bersifat rekreasi dan melibatkan keluarga, dengan kesan yang menyenangkan.
11	<i>Good, up, keep, excit, spirit, pindul, realli, cave, work, friend</i>	Kepuasan dan Pengalaman Wisata	Topik ini menunjukkan tingkat kepuasan pengunjung terhadap

Topik	Keywords Dominan	Label Topik	Interpretasi
			pengalaman wisata secara keseluruhan.
12	<i>Pleas, god, will, swim, cant, si, amen, success, must, possibl</i>	Harapan dan Aktivitas Wisata	Topik ini berkaitan dengan harapan atau keinginan pengunjung terhadap aktivitas wisata yang dilakukan.
13	<i>Open, bro, great, admin, thank, next, again, time, si, good</i>	Interaksi dan Pelayanan Admin	Topik ini menggambarkan interaksi antara pengunjung dan pihak pengelola, khususnya dalam hal pelayanan administrasi atau komunikasi.
14	<i>Ticket, b, thank, about, time, differ, realli, travel, cant, long</i>	Informasi Tiket dan Waktu	Topik ini berkaitan dengan informasi yang dibutuhkan pengunjung mengenai tiket, waktu operasional, dan ketentuan kunjungan.
15	<i>Dm, reserv, document, make, look, happi, refresh, raft, want, upload</i>	Reservasi dan Administrasi Wisata	Topik ini menunjukkan proses reservasi, administrasi, dan pengelolaan dokumen yang berkaitan dengan kegiatan wisata.

5.2. Hasil Analisis Sentimen

Hasil analisis sentimen dilakukan berdasarkan pengolahan data yang telah dilakukan pada BAB IV, khususnya pada bagian 4.4.1 dan 4.4.2. Analisis sentimen dilakukan menggunakan metode berbasis leksikon, yaitu VADER, untuk mengidentifikasi kecenderungan opini pengunjung terhadap wisata Goa Pindul berdasarkan ulasan pada platform *Google Maps* dan *Instagram*.

Tabel 5. 3 Jumlah *label* analisis sentimen ulasan *Google Maps*

Sentimen Label	Jumlah	Persentase
Positif ($\geq 0,05$)	2098	78,8%
Netral ($-0,05 < x < 0,05$)	446	16,8%
Negatif ($\leq -0,5$)	117	4,4%

Berdasarkan hasil analisis sentimen pada data *Google Maps*, diketahui bahwa mayoritas ulasan pengunjung menunjukkan kecenderungan sentimen positif. Hal ini terlihat dari banyaknya nilai *compound score* yang berada di atas ambang batas 0,05, yang mengindikasikan bahwa pengunjung secara umum merasa puas terhadap pengalaman wisata yang ditawarkan. Sentimen positif tersebut umumnya berkaitan dengan pengalaman wisata yang menyenangkan, keindahan alam, serta kualitas pelayanan pemandu wisata yang dinilai ramah dan informatif. Temuan ini menunjukkan bahwa Goa Pindul memiliki daya tarik utama pada aspek pengalaman wisata berbasis alam yang mampu memberikan kesan positif bagi pengunjung.

Namun demikian, hasil analisis juga menunjukkan adanya sentimen negatif meskipun dalam jumlah yang lebih kecil. Sentimen negatif pada ulasan *Google Maps* umumnya berkaitan dengan permasalahan seperti harga, keberadaan calo, serta aspek pengelolaan wisata yang belum optimal. Hal ini menunjukkan bahwa masih terdapat beberapa aspek pelayanan yang belum sepenuhnya memenuhi ekspektasi pengunjung. Selain itu, terdapat pula ulasan dengan sentimen netral yang umumnya berisi pertanyaan atau informasi tanpa menunjukkan opini yang jelas.

Tabel 5. 4 Jumlah *label* sentimen ulasan *Instagram*

Sentimen Label	Jumlah	Persentase
Positif ($\geq 0,05$)	425	49,4%
Netral ($-0,05 < x < 0,05$)	365	42,4%
Negatif ($\leq -0,5$)	71	8,2%

Sementara itu, pada data Instagram, hasil analisis sentimen juga menunjukkan dominasi sentimen positif. Hal ini dapat dilihat dari banyaknya komentar yang mengandung ekspresi apresiasi seperti “*fun*”, “*amazing*”, dan “*cool*”, yang mencerminkan pengalaman positif pengunjung terhadap wisata Goa Pindul. Karakteristik komentar di *Instagram* cenderung lebih singkat dan ekspresif, serta sering kali disertai dengan emoji atau ungkapan informal, sehingga memperkuat kecenderungan sentimen positif dalam platform tersebut.

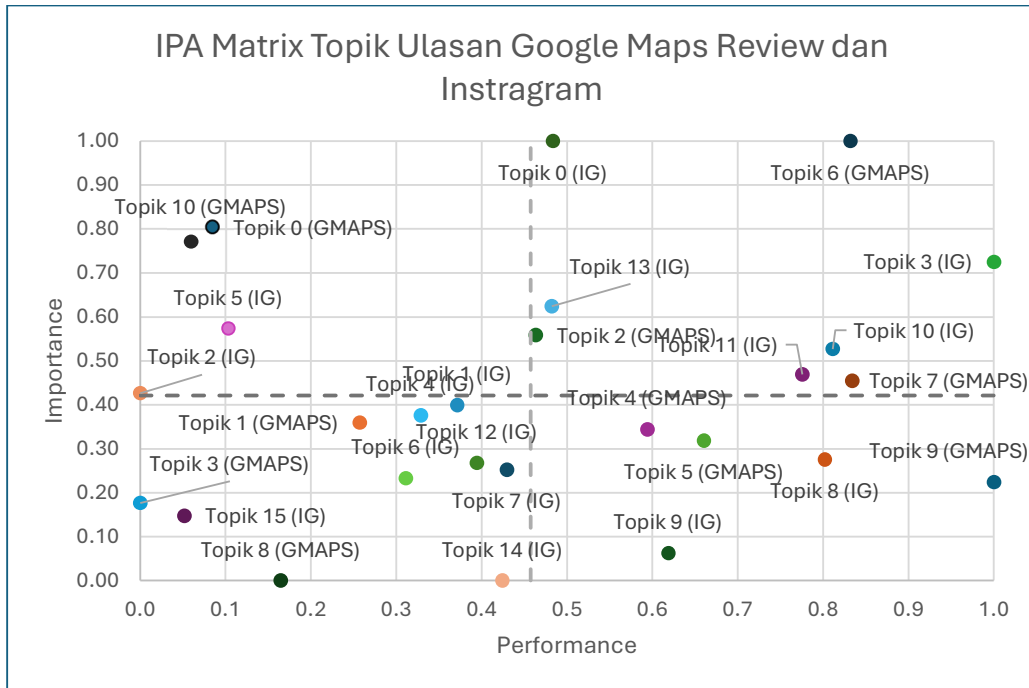
Meskipun demikian, terdapat pula sentimen negatif pada komentar *Instagram*, yang umumnya berkaitan dengan pertanyaan atau kekhawatiran terkait keamanan, usia pengunjung, serta kondisi tertentu dalam aktivitas wisata. Selain itu, beberapa komentar juga menunjukkan kerinduan atau ekspektasi yang belum terpenuhi, yang dalam analisis VADER dapat terklasifikasi sebagai sentimen negatif karena mengandung nuansa emosional tertentu.

Hasil analisis ini sejalan dengan konsep analisis sentimen berbasis leksikon seperti VADER yang efektif digunakan untuk mengolah teks pendek dan informal pada media sosial. Metode ini mampu menangkap polaritas sentimen berdasarkan skor *compound* yang merepresentasikan kecenderungan opini secara keseluruhan. Dengan demikian, penggunaan VADER dalam penelitian ini dinilai sesuai dengan karakteristik data yang berasal dari platform digital.

Secara keseluruhan, hasil analisis sentimen menunjukkan bahwa persepsi pengunjung terhadap wisata Goa Pindul cenderung positif, yang mengindikasikan tingkat kepuasan yang relatif tinggi. Namun, keberadaan sentimen negatif pada beberapa aspek menunjukkan adanya peluang perbaikan, khususnya dalam hal pengelolaan layanan dan pengalaman wisata. Oleh karena itu, hasil analisis sentimen ini menjadi dasar penting dalam menentukan prioritas perbaikan pada tahap analisis selanjutnya menggunakan metode *Importance Performance Analysis* (IPA).

5.3. Hasil *Importance-Performance Analysis* (IPA) Matrix

Pada sub bab ini disajikan hasil pemetaan atribut berdasarkan metode *Importance-Performance Analysis* (IPA) dalam bentuk diagram kartesius. Pemetaan dilakukan dengan menggunakan dua dimensi utama, yaitu tingkat kepentingan (*importance*) sebagai sumbu vertikal (Y) dan tingkat kinerja (*performance*) sebagai sumbu horizontal (X). Nilai rata-rata dari kedua dimensi digunakan sebagai titik pembagi untuk membentuk empat kuadran analisis.



Gambar 5. 1 IPA Matrix Ulasan Google Maps Review dan Instagram

5.3.1. Google Maps IPA Matrix

Berdasarkan hasil visualisasi IPA Matrix pada data *Google Maps*, seluruh topik hasil pemodelan LDA terdistribusi ke dalam empat kuadran dengan karakteristik yang berbeda. Setiap kuadran merepresentasikan kondisi persepsi pengunjung terhadap atribut wisata Goa Pindul.

1. Kuadran I (*Concentrate Here*)

Kuadran ini menunjukkan atribut dengan tingkat kepentingan tinggi namun memiliki kinerja yang relatif rendah, sehingga menjadi prioritas utama dalam evaluasi. Topik yang termasuk dalam kuadran ini adalah:

a. Topik 2 (Kebersihan Fasilitas dan Keindahan Alam)

Keberadaan topik ini dalam Kuadran I menunjukkan bahwa aspek utama yang menjadi daya tarik wisata, yaitu kebersihan fasilitas dan keindahan alam, namun kinerjanya belum sepenuhnya memenuhi harapan. Hal ini mengindikasikan adanya kesenjangan antara ekspektasi dan pengalaman aktual yang dirasakan pengunjung pada topik ini. Dengan demikian, atribut pada kuadran ini menjadi fokus utama dalam evaluasi kualitas layanan wisata Goa Pindul,

2. Kuadran II (*Keep Up the Good Work*)

Kuadran ini berisi atribut dengan tingkat kepentingan dan kinerja yang sama-sama tinggi, sehingga menunjukkan performa yang sudah sesuai dengan harapan pengunjung.

Topik yang berada pada kuadran ini antara lain:

- a. Topik 0 (Keseruan Wisata Goa Pindul)
- b. Topik 6 (Keseruan *Tour* dan Keramahan Pemandu)
- c. Topik 7 (Liburan di Goa Pindul)
- d. Topik 10 (Kepuasan Layanan untuk Keluarga)

Keberadaan topik-topik tersebut menunjukkan bahwa aspek keseruan wisata, keramahan pemandu, liburan, hingga kepuasan layanan untuk keluarga telah mampu memenuhi ekspektasi pengunjung. Atribut pada kuadran ini mencerminkan kekuatan utama destinasi yang perlu dipertahankan kualitasnya.

3. Kuadran III (*Low Priority*)

Kuadran ini mencakup atribut dengan tingkat kepentingan dan kinerja yang sama-sama rendah, sehingga tidak menjadi prioritas utama dalam pengembangan. Topik yang termasuk dalam kuadran ini meliputi:

- a. Topik 1 (Aktivitas *Rafting*)
- b. Topik 8 (Pengalaman Pertama ke Goa Pindul)

Posisi topik-topik tersebut menunjukkan bahwa aspek ini tidak terlalu menjadi perhatian utama bagi pengunjung, serta kinerjanya juga tidak terlalu menonjol dibandingkan atribut lainnya.

4. Kuadran IV (*Possible Overkill*)

Kuadran ini menunjukkan atribut dengan tingkat kinerja tinggi namun tingkat kepentingannya relatif rendah, yang mengindikasikan adanya potensi kelebihan dalam alokasi sumber daya. Topik yang berada di kuadran ini antara lain:

- a. Topik 3 (Kepuasan dan Pengalaman Wisata)
- b. Topik 4 (Eksplorasi Wisata *Cave Tubing*)
- c. Topik 5 (Pengalaman Petualangan Wisata)
- d. Topik 9 (Pelayanan Wisata)

Keberadaan topik-topik tersebut menunjukkan bahwa aspek-aspek tersebut telah memiliki kinerja yang baik, namun tidak sepenuhnya dianggap sebagai prioritas utama oleh pengunjung.

Secara keseluruhan, hasil IPA *Matrix* menunjukkan bahwa meskipun sebagian besar fasilitas pelayanan wisata Goa Pindul telah memiliki kinerja yang baik, masih terdapat aspek penting yang perlu menjadi perhatian utama, khususnya pada Kuadran I. Distribusi topik yang tersebar di seluruh kuadran menunjukkan bahwa persepsi pengunjung terhadap wisata Goa Pindul bersifat beragam, baik dari sisi kepentingan maupun kinerja.

5.3.2. Instagram IPA *Matrix*

Berdasarkan hasil visualisasi IPA *Matrix* pada data *Google Maps*, seluruh topik hasil pemodelan LDA terdistribusi ke dalam empat kuadran dengan karakteristik yang berbeda. Setiap kuadran merepresentasikan kondisi persepsi pengunjung terhadap atribut wisata Goa Pindul.

1. Kuadran I (*Concentrate Here*)

Kuadran ini menunjukkan atribut dengan tingkat kepentingan tinggi namun memiliki kinerja yang relatif rendah, sehingga menjadi prioritas utama dalam evaluasi. Topik yang termasuk dalam kuadran ini antara lain:

- a. Topik 10 (Pengalaman Keluarga dan Hiburan)
- b. Topik 11 (Kepuasan Wisatawan Goa Pindul)
- c. Topik 13 (Interaksi dan Pelayanan Admin)

Keberadaan topik-topik ini menunjukkan bahwa aspek utama yang berkaitan dengan pengalaman keluarga, kepuasan wisatawan, dan interaksi serta pelayanan *admin* merupakan hal yang sangat penting bagi pengunjung. Namun, kinerja pada aspek tersebut belum sepenuhnya memenuhi harapan. Hal ini mengindikasikan adanya kesenjangan antara ekspektasi dan pengalaman aktual, terutama pada topik-topik tersebut pada pengalaman wisata.

2. Kuadran 2 (*Keep Up the Good Work*)

Kuadran ini berisi atribut dengan tingkat kepentingan dan kinerja yang sama-sama tinggi. Topik yang berada pada kuadran ini antara lain:

- a. Topik 0 (Keamanan Anak dan Batas Usia)
- b. Topik 3 (Aktivitas Wisata Keluarga)
- c. Topik 5 (Harapan Wisatawan terhadap Goa Pindul)

Topik-topik tersebut menunjukkan bahwa aspek keamanan, pelayanan pengelola, serta pengalaman wisata berbasis keluarga telah memiliki kinerja yang baik dan sesuai

dengan harapan pengunjung. Hal ini mencerminkan kekuatan utama wisata Goa Pindul dalam memberikan pengalaman yang aman dan ramah bagi wisata keluarga.

3. Kuadran III (*Low Priority*)

Kuadran ini mencakup atribut dengan tingkat kepentingan dan kinerja yang sama-sama rendah. Topik yang termasuk dalam kuadran ini antara lain:

- a. Topik 1 (Aktivitas Seru dan Rafting)
- b. Topik 2 (Tiket dan Akses Masuk)
- c. Topik 7 (Keluhan dan Mitigasi Wisata)
- d. Topik 8 (Ucapan Terima Kasih Pengunjung)
- e. Topik 12 (Harapan dan Aktivitas Wisata)
- f. Topik 14 (Informasi Tiket dan Waktu)
- g. Topik 15 (Reservasi dan Administrasi Wisata)

Topik-topik tersebut menunjukkan bahwa aspek aspek yang ada tidak menjadi perhatian utama dalam konteks komentar Instagram. Selain itu, kinerja pada aspek tersebut juga tidak terlalu menonjol dibandingkan atribut lainnya.

4. Kuadran IV (*Possible Overkill*)

Kuadran ini menunjukkan atribut dengan tingkat kinerja tinggi namun tingkat kepentingannya relatif rendah. Topik yang berada pada kuadran ini antara lain:

- a. Topik 4 (Kekhawatiran dan Risiko Wisata)
- b. Topik 6 (Keindahan Sungai dan Wisata Air)

Keberadaan topik-topik ini menunjukkan bahwa kekhawatiran terhadap resiko wisata dan keindahan wisata air Goa Pindul memiliki kinerja yang baik dalam mencerminkan pengalaman positif. Namun, aspek tersebut tidak dianggap sebagai prioritas utama dalam menentukan kualitas layanan wisata secara keseluruhan.

Secara keseluruhan, hasil IPA Matrix pada data Instagram menunjukkan bahwa terdapat beberapa atribut penting yang masih belum optimal, terutama pada kebersihan fasilitas, keindahan alam, dan kepuasan wisatawan yang menjadi aspek utama pengunjung. Di sisi lain, aspek keamanan, pelayanan, dan pengalaman wisata keluarga telah menunjukkan kinerja yang baik dan menjadi kekuatan utama destinasi.

5.4. Rekomendasi Strategi Pengembangan Wisata Goa Pindul

Berdasarkan hasil pemodelan topik menggunakan *Latent Dirichlet Allocation* (LDA), analisis sentimen menggunakan VADER, dan pemetaan *Importance Performance Analysis* (IPA),

diperoleh empat topik yang berada pada Kuadran I (*Concentrate Here*). Topik-topik tersebut memiliki tingkat kepentingan yang tinggi namun tingkat kinerja yang relatif rendah sehingga menjadi prioritas utama perbaikan. Hasil analisis tersebut diterjemahkan menjadi usulan rancangan sistem yang dapat digunakan oleh pengelola Goa Pindul sebagai dasar pengembangan layanan wisata. Berikut merupakan beberapa rekomendasi strategi yang dapat dijadikan sebagai dasar perbaikan bagi pengelola wisata Goa Pindul.

5.4.1. Perancangan Sistem Audit Kebersihan dan Pengendalian Lingkungan Wisata

Rekomendasi ini disusun berdasarkan Topik 2 *Google Maps* yaitu Kebersihan Fasilitas dan Keindahan Alam. Hasil IPA menunjukkan bahwa aspek kebersihan dan kualitas lingkungan wisata merupakan atribut yang penting bagi pengunjung, namun kinerjanya belum memenuhi harapan wisatawan.

Sebagai upaya perbaikan, pengelola Goa Pindul disarankan merancang sistem audit kebersihan yang terstandarisasi. Sistem ini bertujuan untuk memastikan kondisi fasilitas dan lingkungan wisata selalu berada pada standar yang telah ditetapkan. Rancangan sistem meliputi penyusunan *Standard Operating Procedure* (SOP) kebersihan, *checklist* inspeksi harian, jadwal audit rutin, serta indikator kinerja utama (*Key Performance Indicator/KPI*) kebersihan.

Indikator yang dapat digunakan antara lain tingkat kebersihan toilet, area parkir, area tunggu wisatawan, jalur masuk wisata, serta kondisi kebersihan sungai dan area sekitar Goa Pindul. Hasil evaluasi kemudian direkapitulasi dalam bentuk skor kebersihan yang digunakan sebagai dasar tindakan perbaikan berkelanjutan. Dengan adanya sistem audit tersebut, pengelola dapat melakukan pengendalian kualitas layanan secara lebih sistematis dan terukur.

Tabel 5. 5 Sistem Audit Kebersihan

Fasilitas	Total Item	Standar	Jumlah Item Sesuai Standar	Skor Kebersihan
Toilet	4	Tidak ada bau		
		Tidak ada sampah		
		Wadah air dalam keadaan bersih		
		Sirkulasi air mengalir dengan baik		
		Kondisi toilet dengan penerangan baik		

Fasilitas	Total Item	Standar	Jumlah Item Sesuai Standar	Skor Kebersihan
Area Parkir	1	Tidak ada sampah plastik		
		Tidak ada pohon tumbang		
Area Tunggu	2	Tidak ada sampah		
		Tidak ada genangan		
Jalur Masuk Wisata	1	Tidak ada sampah		
		Tidak ada jalan berlubang		
Sungai	1	Tidak ada limbah yang terlihat		

$$\text{Skor Kebersihan} = \frac{\text{Jumlah item sesuai standar}}{\text{Total item}} \times 100\%$$

Skor kebersihan yang didapatkan merepresentasikan tingkat kelayakan fasilitas dalam segi kebersihan. Pengelola wisata Goa Pindul dapat menetapkan target skor kebersihan yang sesuai dengan kemampuan maksimal pengelola, misalnya dengan target skor kebersihan sebesar $\geq 90\%$ untuk tiap-tiap fasilitas yang tersedia.

5.4.2. Perancangan Customer Journey Wisata Keluarga

Rekomendasi ini disusun berdasarkan Topik 10 Instagram yaitu Pengalaman Keluarga dan Hiburan. Hasil IPA menunjukkan bahwa wisatawan memiliki harapan tinggi terhadap pengalaman wisata keluarga, namun pengalaman yang diterima masih belum optimal. Untuk mengatasi permasalahan tersebut, pengelola disarankan merancang *Customer Journey* wisata keluarga. *Customer Journey* merupakan alat pemetaan proses layanan yang menggambarkan seluruh interaksi antara wisatawan dan penyedia layanan mulai dari kedatangan hingga wisata selesai.

Pada penelitian ini, customer journey dibagi menjadi lima tahapan utama, yaitu perencanaan kunjungan (*planning*), kedatangan (*arrival*), pelaksanaan wisata (*experience*), penyelesaian kunjungan (*departure*), dan pascakunjungan (*post-visit*). Setiap tahapan menggambarkan aktivitas utama wisatawan, kebutuhan pelanggan, serta peluang perbaikan yang dapat dilakukan oleh pengelola.

Tabel 5. 6 Perancangan *Customer Journey* Wisata Keluarga

Tahapan	Aktivitas Wisata	Touchpoint	Potensi Permasalahan	Usulan Perbaikan
<i>Planning</i>	Mencari informasi mengenai Goa Pindul	<i>Instagram, Google Maps</i>	Informasi kurang lengkap mengenai fasilitas keluarga	Menyediakan informasi fasilitas keluarga, paket wisata, usia minimum, dan estimasi durasi wisata
<i>Arrival</i>	Datang ke lokasi dan melakukan registrasi	Area parkir, loket tiket	Antrean panjang, area tunggu kurang nyaman bagi keluarga	Menambah area tunggu, sistem antrean yang lebih teratur, dan jalur khusus keluarga
<i>Experience</i>	Mengikuti kegiatan cave tubing bersama keluarga	Pemandu wisata, area sungai	Aktivitas pendukung keluarga masih terbatas	Menambah area bermain anak, spot foto keluarga, dan aktivitas edukasi ringan
<i>Departure</i>	Membersihkan diri dan meninggalkan lokasi	Toilet, ruang bilas, area keluar	Kepadatan ruang bilas setelah wisata	Menambah kapasitas ruang bilas dan memperbaiki alur keluar pengunjung
<i>Post-Visit</i>	Memberikan ulasan dan membagikan pengalaman	<i>Instagram, Google Maps</i>	Tidak ada tindak lanjut terhadap masukan wisatawan	<i>Monitoring</i> ulasan secara berkala sebagai bahan evaluasi layanan

Berdasarkan customer journey yang telah dirancang, diketahui bahwa pengalaman wisata keluarga tidak hanya dipengaruhi oleh aktivitas utama cave tubing, tetapi juga oleh kualitas

pelayanan pada setiap tahapan perjalanan wisatawan. Tahap *planning* menjadi awal terbentuknya ekspektasi pengunjung melalui informasi yang tersedia pada media *digital*. Oleh karena itu, penyediaan informasi yang lengkap mengenai fasilitas keluarga, batas usia, harga, dan paket wisata menjadi faktor penting dalam membangun persepsi awal wisatawan.

Pada tahap *arrival*, kenyamanan wisatawan dipengaruhi oleh proses registrasi, kondisi area parkir, dan fasilitas ruang tunggu. Hasil pemetaan menunjukkan bahwa keluarga dengan anak-anak membutuhkan area tunggu yang lebih nyaman serta sistem antrean yang lebih terorganisasi agar pengalaman sebelum wisata tetap positif.

Tahap *experience* merupakan inti perjalanan wisata karena wisatawan berinteraksi langsung dengan atraksi utama Goa Pindul. Berdasarkan hasil analisis, pengelola dapat meningkatkan pengalaman wisata dengan menyediakan fasilitas yang lebih ramah keluarga, seperti area bermain anak, spot dokumentasi keluarga, serta aktivitas edukatif mengenai lingkungan Goa Pindul. Pendekatan ini diharapkan mampu meningkatkan nilai pengalaman (*customer experience*) tanpa mengubah atraksi utama wisata.

Pada tahap *departure*, perhatian perlu diberikan pada fasilitas pendukung seperti ruang bilas, toilet, dan alur keluar pengunjung. Pengelolaan kapasitas fasilitas tersebut dapat mengurangi kepadatan setelah aktivitas wisata sehingga kenyamanan wisatawan tetap terjaga hingga akhir kunjungan.

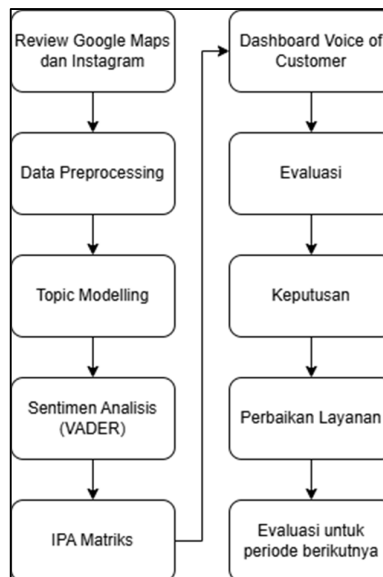
Selanjutnya, pada tahap *post-visit*, ulasan yang diberikan wisatawan melalui *Google Maps* maupun *Instagram* dapat dimanfaatkan sebagai sumber *Voice of Customer* untuk mengevaluasi kualitas layanan secara berkelanjutan. Dengan demikian, *customer journey* tidak hanya digunakan sebagai alat pemetaan pengalaman wisatawan, tetapi juga menjadi dasar dalam merancang perbaikan layanan yang berorientasi pada kebutuhan pelanggan.

Melalui *Customer Journey*, pengelola dapat mengidentifikasi titik-titik kritis yang berpotensi menimbulkan ketidakpuasan wisatawan, seperti waktu tunggu yang terlalu lama, keterbatasan fasilitas keluarga, atau kurangnya aktivitas pendukung bagi anak-anak. Hasil identifikasi tersebut dapat digunakan sebagai dasar pengembangan fasilitas ramah keluarga, penyediaan area bermain anak, area istirahat keluarga, serta pengembangan paket wisata keluarga yang lebih terintegrasi.

5.4.3. Perancangan Dashboard Voice of Customer (VOC)

Rekomendasi ini disusun berdasarkan Topik 11 Instagram yaitu Kepuasan Wisatawan Goa Pindul. Hasil analisis menunjukkan bahwa tingkat kepuasan wisatawan menjadi aspek yang sangat penting, namun masih terdapat kesenjangan antara harapan dan pengalaman aktual pengunjung. Sebagai solusi, pengelola disarankan merancang sistem *Voice of Customer* (VoC) berbasis *dashboard digital* yang berfungsi untuk memantau persepsi wisatawan secara berkelanjutan. *Dashboard* tersebut dapat mengintegrasikan data ulasan dari *Google Maps* maupun *Instagram* untuk menghasilkan informasi mengenai tren sentimen, topik yang paling sering dibahas, serta keluhan utama wisatawan.

Dashboard Voice of Customer (VoC) yang diusulkan memiliki alur kerja sebagai berikut:

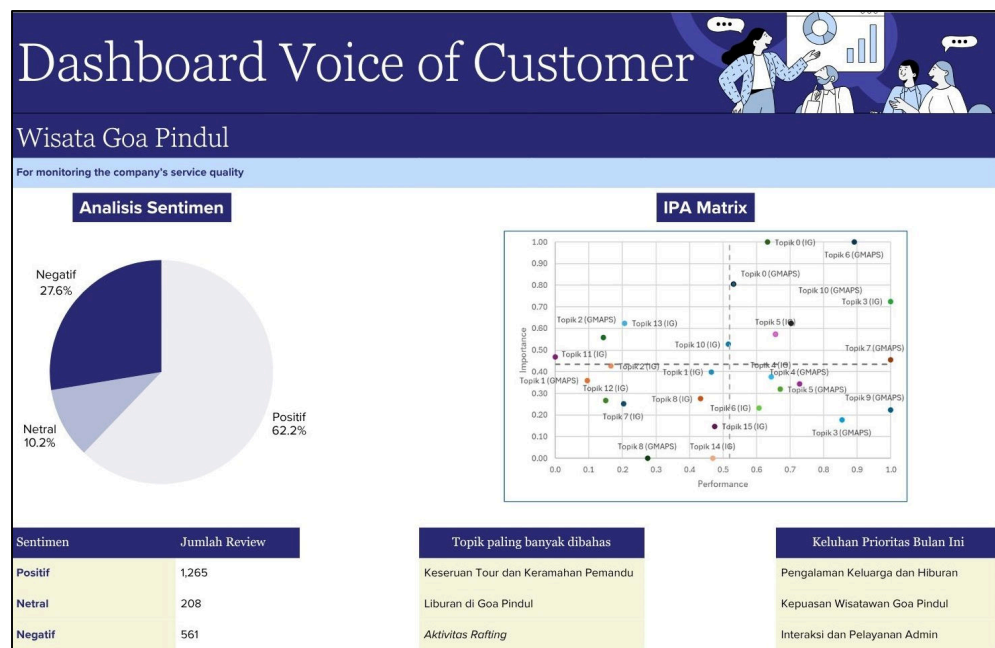


Gambar 5. 2 Alur *Dashboard VoC*

Alur tersebut menunjukkan bahwa *dashboard* tidak melakukan analisis secara langsung, melainkan menerima hasil pengolahan data yang telah dilakukan pada penelitian ini. *Dashboard* kemudian menyajikan informasi dalam bentuk visual sehingga memudahkan pengelola melakukan monitoring kualitas layanan secara berkelanjutan.

Dashboard yang diusulkan terdiri atas lima komponen utama, yaitu jumlah *review* yang menunjukkan jumlah ulasan yang masuk dari masing-masing *platform*, distribusi sentimen untuk memantau perubahan kepuasan wisatawan setiap periode, topik dominan untuk mengetahui isu yang paling sering dibahas wisatawan, IPA Matrix untuk menentukan prioritas perbaikan layanan, dan daftar keluhan yang menampilkan keluhan utama yang harus segera

ditindaklanjuti. Berikut merupakan usulan *design dashboard Voice of Customer* bagi pengelola wisata Goa Pindul:



Gambar 5. 3 *Dashboard Voice of Customer*

Informasi yang diperoleh dari *dashboard* dapat digunakan sebagai dasar pengambilan keputusan dalam pengelolaan wisata. Selain itu, *dashboard* juga memungkinkan pengelola melakukan evaluasi layanan secara berkala dan mendeteksi permasalahan yang muncul lebih cepat dibandingkan metode evaluasi konvensional. Implementasi *dashboard VoC* diharapkan mampu mendukung pengambilan keputusan berbasis data serta menciptakan sistem perbaikan layanan yang berkelanjutan.

5.4.4. Perancangan SOP dan Standarisasi Pelayanan Admin

Rekomendasi ini disusun berdasarkan Topik 13 Instagram yaitu Interaksi dan Pelayanan Admin. Hasil IPA menunjukkan bahwa kualitas komunikasi dan pelayanan administrasi merupakan aspek yang penting bagi wisatawan, namun performanya masih belum memenuhi ekspektasi pengunjung. Untuk meningkatkan kualitas pelayanan tersebut, pengelola disarankan merancang *Standard Operating Procedure* (SOP) pelayanan admin yang mencakup alur penanganan pertanyaan, reservasi, keluhan, dan pemberian informasi kepada wisatawan. SOP ini perlu dilengkapi dengan standar waktu respons, standar kualitas pelayanan, serta indikator kinerja pelayanan.

Tabel 5. 7 SOP dan Standarisasi Pelayanan Admin

Standar Operating Procedure (SOP)	
PROSEDUR PELAYANAN ADMINISTRASI WISATA GOA PINDUL	
Tanggal Terbit	Ditetapkan oleh: Manager Goa Pindul
Penjelasan	SOP ini digunakan sebagai pedoman pelayanan admin dalam memberikan informasi, menangani reservasi, serta merespons pertanyaan dan keluhan wisatawan melalui media komunikasi resmi Goa Pindul sehingga pelayanan menjadi lebih cepat, konsisten, dan berkualitas.
Tujuan	Menstandarkan proses pelayanan admin agar seluruh wisatawan memperoleh informasi yang akurat, pelayanan yang responsif, serta penanganan keluhan yang efektif sehingga dapat meningkatkan kepuasan wisatawan.
Kebijakan	Seluruh petugas admin wajib mengikuti prosedur pelayanan yang telah ditetapkan. Seluruh komunikasi dengan wisatawan harus menggunakan informasi resmi Goa Pindul serta didokumentasikan sebagai bahan evaluasi kualitas pelayanan.
Prosedur	
Aktivitas Pekerjaan	Langkah Kerja
Penerimaan Pertanyaan Wisatawan	• Menerima pertanyaan dari <i>Instagram</i>, <i>Whatsapp</i>, atau telepon • Mencatat waktu masuk pertanyaan • Memastikan identitas atau kebutuhan wisatawan
Identifikasi Kebutuhan Wisatawan	• Mengelompokkan pertanyaan menjadi informasi umum, reservasi, atau keluhan

	• Memprioritaskan pelayanan dengan cara memulai dari permintaan yang masuk paling awal apabila terdapat banyak permintaan secara bersamaan.
Pemberian Informasi	• Memberikan informasi mengenai harga tiket, paket wisata, jam operasional, fasilitas, lokasi, dan ketentuan wisata. • Menggunakan informasi berdasarkan <i>database</i> atau FAQ resmi • Menghindari penyampaian informasi yang belum terverifikasi.
Penanganan Reservasi dan Keluhan	• Memproses permintaan reservasi sesuai prosedur • Mencatat setiap keluhan wisatawan • Meneruskan keluhan kepada bagian terkait apabila membutuhkan tindak lanjut • Memantau hingga keluhan dinyatakan selesai.
Konfirmasi Pelayanan	• Memastikan seluruh pertanyaan telah terjawab • Mengonfirmasi kembali kepada wisatawan bahwa pelayanan telah selesai • Mengucapkan terima kasih atas kunjungan atau pertanyaan wisatawan.
Dokumentasi Pelayanan	• Mencatat seluruh aktivitas pelayanan dalam <i>log</i> pelayanan • Menyimpan data sebagai bahan evaluasi dan <i>monitoring</i> kualitas pelayanan
Monitoring dan Evaluasi (KPI)	
Indikator Kerja	Target
<i>Response Time</i>	$\leq 15 \text{ menit}$
<i>Resolution Time</i>	$\leq 1 \times 24 \text{ jam}$
Akurasi informasi yang diberikan	$\geq 95\%$
Persentasi keluhan yang terselesaikan	$\geq 90\%$
Tingkat kepuasan pelayanan admin	$\geq 90\%$

Kelengkapan dokumentasi pelayanan	100%
Keterangan: 1. SOP ini dilaksanakan oleh seluruh petugas admin yang bertugas melayani wisatawan melalui media komunikasi resmi Goa Pindul. 2. Seluruh pertanyaan, reservasi, dan keluhan wisatawan wajib dicatat pada log pelayanan sebagai bahan evaluasi. 3. Informasi yang diberikan kepada wisatawan harus mengacu pada informasi resmi mengenai harga tiket, jam operasional, paket wisata, dan ketentuan lainnya. 4. Keluhan yang tidak dapat diselesaikan oleh admin wajib diteruskan kepada bagian operasional atau pengelola untuk memperoleh tindak lanjut. 5. Evaluasi terhadap pelaksanaan SOP dilakukan secara berkala menggunakan indikator kinerja pelayanan (KPI) sebagai dasar perbaikan layanan. 6. Hasil evaluasi digunakan sebagai masukan dalam peningkatan kualitas pelayanan dan pengambilan keputusan oleh pengelola Goa Pindul.	

Standarisasi proses pelayanan akan membantu menciptakan konsistensi layanan serta meningkatkan kecepatan dan akurasi informasi yang diterima wisatawan. Dengan adanya SOP pelayanan admin yang terstruktur, kualitas interaksi antara wisatawan dan pengelola dapat meningkat sehingga mampu mendukung terciptanya pengalaman wisata yang lebih baik.

BAB VI

PENUTUP

6.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan melalui tahapan pengolahan data, analisis sentimen, pemodelan topik menggunakan *Latent Dirichlet Allocation* (LDA), serta *Importance Performance Analysis* (IPA), maka diperoleh beberapa kesimpulan yang menjawab tujuan penelitian sebagai berikut:

1. Hasil pemodelan topik menggunakan metode LDA menunjukkan bahwa ulasan pengunjung wisata Goa Pindul pada *platform Google Maps* dan *Instagram* menghasilkan beberapa topik utama yang merepresentasikan persepsi pengunjung. Pada data *Google Maps*, terdapat sebelas topik yang terbentuk. Topik dengan nilai *importance* tertinggi adalah Topik 6 Keseruan *Tour* dan Keramahan Pemandu (1), Topik 0 Keseruan Wisata Goa Pindul (0,804), dan Topik 10 Kepuasan Layanan untuk Keluarga (0,779). Sementara itu, pada data *Instagram*, topik yang muncul lebih beragam dengan total topik yang terbentuk adalah sebanyak 16 topik. Topik dengan nilai *importance* tertinggi adalah Topik 0 Keamanan Anak dan Batas Usia (1), Topik 3 Aktivitas Wisata Keluarga (0,724), dan Topik 13 Reservasi dan Administrasi Wisata (0,624). Hal ini menunjukkan bahwa LDA mampu mengidentifikasi struktur pembahasan utama dalam ulasan pengunjung, di mana *Google Maps* lebih bersifat evaluatif terhadap pengalaman, sedangkan *Instagram* lebih mencerminkan interaksi dan ekspresi pengalaman secara singkat.
2. Berdasarkan hasil analisis sentimen menggunakan metode VADER, kecenderungan sentimen pengunjung terhadap wisata Goa Pindul didominasi oleh sentimen positif. Hal ini menunjukkan bahwa secara umum pengunjung merasa puas terhadap pengalaman wisata yang diperoleh, terutama pada aspek keunikan wisata *cave tubing*, suasana alam, dan pengalaman rekreasi yang ditawarkan. Meskipun demikian, masih ditemukan sentimen negatif pada beberapa ulasan. Topik pada *Google Maps* yang memiliki nilai *performance* tertinggi adalah pada Topik 9 Pelayanan Wisata (1) dan Topik 7 Musim Libur di Goa Pindul (0,834). Sedangkan topik pada *Instagram* yang memiliki nilai *performance* tertinggi adalah pada Topik 3 Aktivitas Wisata Keluarga (1) dan Topik 10 Pengalaman Keluarga dan Hiburan (0,811).

3. Berdasarkan hasil *Importance Performance Analysis* (IPA), diperoleh atribut-atribut yang menjadi prioritas perbaikan pada Kuadran I, yaitu atribut dengan tingkat kepentingan tinggi namun memiliki kinerja yang rendah. Pada data *Google Maps*, atribut prioritas meliputi keseruan wisata goa pindul serta kepuasan layanan untuk keluarga. Sedangkan pada data Instagram meliputi aspek tiket dan akses masuk, serta harapan wisatawan terhadap goa pindul. Hasil ini menunjukkan bahwa aspek keseruan wisata goa pindul justru menjadi area yang perlu ditingkatkan. Selain itu, temuan dari analisis sentimen juga mengindikasikan adanya keluhan terkait transparansi harga, praktik percaloan, serta pengelolaan layanan, yang memperkuat hasil pemetaan IPA bahwa aspek operasional dan pengalaman wisata masih perlu diperbaiki.
4. Berdasarkan prioritas permasalahan yang diperoleh, strategi pengembangan wisata Goa Pindul difokuskan pada empat aspek utama, yaitu peningkatan kebersihan fasilitas dan keindahan alam, pengalaman keluarga dan hiburan, kepuasan wisatawan goa pindul, serta interaksi dan pelayanan admin. Strategi yang direkomendasikan meliputi perancangan sistem audit kebersihan dan fasilitas, perancangan *customer journey* wisata keluarga, perancangan *dashboard voice of customer*, serta perancangan SOP dan standarisasi pelayanan admin. Dengan mengacu pada hasil analisis berbasis data ulasan digital, strategi ini diharapkan mampu meningkatkan kepuasan pengunjung serta memperkuat daya saing wisata Goa Pindul secara berkelanjutan.

6.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan baik bagi pengelola wisata Goa Pindul maupun bagi penelitian selanjutnya:

1. Saran bagi Pengelola Wisata Goa Pindul

Pengelola wisata Goa Pindul dapat menggunakan usulan rekomendasi strategi pada penelitian ini sebagai dasar pengambil keputusan, akan tetapi pengelola wisata disarankan untuk melakukan penyesuaian kembali terkait standar dan prosedur pada usulan rekomendasi ini dengan kapabilitas sumber daya yang relevan. Pengelola wisata Goa Pindul juga disarankan dapat memanfaatkan data komentar yang tersedia di *platform digital* untuk dapat melakukan perbaikan berkelanjutan dengan berbasis persepsi para pengunjung. Selain itu, Penelitian ini diharapkan dapat menjadi dasar dalam pengambilan keputusan untuk peningkatan kualitas layanan wisata Goa Pindul.

2. Saran bagi Penelitian Berikutnya

Penelitian selanjutnya disarankan untuk mengembangkan metode analisis yang lebih komprehensif, seperti penggunaan pendekatan *machine learning* atau *deep learning* dalam analisis sentimen guna meningkatkan akurasi dan pemahaman konteks bahasa. Selain itu, penelitian berikutnya juga dapat memperluas sumber data dengan mengintegrasikan lebih banyak platform digital, serta mengombinasikan pendekatan kuantitatif dengan metode kualitatif seperti wawancara atau survei langsung. Dengan demikian, hasil penelitian diharapkan dapat memberikan gambaran yang lebih mendalam dan komprehensif terhadap persepsi wisatawan.

DAFTAR PUSTAKA

- Adhe, D., Rachman, C., Goejantoro, R., Deny, F., & Amijaya, T. (2020). Implementasi Text Mining Pengelompokan Dokumen Skripsi Menggunakan Metode K-Means Clustering Implementation Of Text Mining For Grouping Thesis Documents Using K-Means Clustering. *Jurnal EKSPONENSIAL*, 11(2).
- Al Farizal, R., & Nugroho, W. A. A. (2025). *Analisis Sentimen Berbasis Aspek Pada Ulasan Google Maps (Aspect-Based Sentiment Analysis on Google Maps Reviews: A Case Study of Tourism Sector Recovery Strategies Post Covid-19 Pandemic in West Papua Province)*.
- Alamsyah, A., Shafira, S., & Yudhistira, A. (2021). *Summarizing Online Conversation of Indonesia Tourism Industry using Network Text Analysis*.
- Ali Laudza, H., Aji Putra, J. T., Attsaqif, M. D., Setiawan, I., & Sahal Annabil, M. (2025). Sentiment Analysis and Topic Modelling of Tourist Reviews on Heritage Destinations of Lawang Sewu. *International Journal of Advances in Scientific Research and Engineering*, 11(09), 38–48. <https://doi.org/10.31695/IJASRE.2025.9.3>
- Ali, T., Marc, B., Omar, B., Soulaïmane, K., & Larbi, S. (2021). Analyzing tourism reviews using an LDA topic-based sentiment analysis approach. *Tourism Management Perspectives*, 40. <https://doi.org/10.1016/j.tmp.2021.100892>
- Ariansyah, A., & Indahyanti, U. (2024). Fitur Ekstraksi pada Pemodelan Topik Menggunakan Metode Latent Dirichlet Allocation pada Peristiwa Kebocoran Data. *Indonesian Journal of Applied Technology*, 1(2), 24. <https://doi.org/10.47134/ijat.v1i2.3041>
- Arianto, D., & Budi, I. (2020). *Aspect-based Sentiment Analysis on Indonesia's Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples)*. <https://maps.google.com/localguides>
- Asyraini, S. (2023). Penerapan Metode Importance Performance Analysis Terhadap Kepuasan Konsumen pada Pelayanan Jasa Asrama Universitas AmirHamzah. *JEBMA: Jurnal Ekonomi Bisnis, Manajemen Dan Akuntansi*. <https://doi.org/10.47709/jebma.v3n2.2777>
- Badan Pusat Statistik. (2026). *Statistik Wisatawan Nusantara*. <https://www.bps.go.id/id>
- Baskara, R. (2022). IMPLEMENTASI WEB SCRAPING PADA MEDIA SOSIAL INSTAGRAM. In *Universitas Islam Indonesia*.
- Blanco-Moreno, S., González-Fernández, A. M., Muñoz-Gallego, P. A., & Casaló, L. V. (2024). Understanding engagement with Instagram posts about tourism destinations. *Journal of Destination Marketing and Management*, 34. <https://doi.org/10.1016/j.jdmm.2024.100948>
- Bratić, M., Pavlović, D., Kovačić, S., Pivac, T., Marić Stanković, A., Vujičić, M. D., & Anđelković, Ž. (2024). An IPA Analysis of Tourist Perception and Satisfaction with Nisville Jazz Festival Service Quality. *Sustainability (Switzerland)*, 16(22). <https://doi.org/10.3390/su16229616>
- Damiana. (2025). ANALISIS EFEKTIFITAS MEDIA SOSIAL INSTAGRAM SEBAGAI SARANA PROMOSI PADA KANTOR POS INDONESIA CABANG SINTANG. *Jurnal Universitas Kapuas Sintang*.
- Deolika, A., & Taufiq Luthfi, E. (2019). ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING. *Jurnal Teknologi Informasi*, 3(2).
- Dian Anggraini, L., Deoranto, P., & Morita Ikasari, D. (2015). ANALISIS PERSEPSI KONSUMEN MENGGUNAKAN METODE IMPORTANCE PERFORMANCE ANALYSIS DAN CUSTOMER SATISFACTION INDEX THE ANALYSIS OF CONSUMER PERCEPTION USED IMPORTANCE PERFORMANCE ANALYSIS

- METHOD AND CUSTOMER SATISFACTION INDEX. *Industria: Jurnal Teknologi Dan Manajemen Agroindustri*, 4.
- Egger, R., & Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Frontiers in Sociology*, 7. <https://doi.org/10.3389/fsoc.2022.886498>
- Fauzi, R., Nasution, H. N., Hastini, F., Zainy, A., & Lumban Tobing, Y. R. (2022). PEGUNAAN MEDIA ADOBE FLASH TERHADAP HASIL BELAJAR SISWA SMKN 1 TANTOM ANGKOLA. *JURNAL EDUCATION AND DEVELOPMENT*, 11(1), 437–442. <https://doi.org/10.37081/ed.v11i1.2687>
- Habets, O., Van der Heijden, B., Ramzy, O., Stoffers, J., & Peters, P. (2021). Employable through social media: An intervention study. *Sustainability (Switzerland)*, 13(9). <https://doi.org/10.3390/su13095093>
- Hapsari, T. S., & Nataliani, Y. (2024). TEXT MINING WITH LATENT DIRICHLET ALLOCATION FOR ANALYZING PUBLIC COMMENTS ON THE M-PASSPORT APPLICATION. *Jurnal Teknik Informatika (Jutif)*, 5(4), 1093–1101. <https://doi.org/10.52436/1.jutif.2024.5.4.1929>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 216–225. <http://sentiment.net/>
- Ipmawati, J., Saifulloh, S., & Kusnawi, K. (2024). Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 247–256. <https://doi.org/10.57152/malcom.v4i1.1066>
- Karo Karo, I. M., Karo Karo, J. A., Djasmayena, S., & Wahyudi, R. (2025). Sentiment Analysis of Tourist Reviews at Waterfront City Pangururan Using Naive Bayes and TF-IDF Algorithm. *Journal of Software Engineering, Information and Communication Technology (SEICT)*, 6(1), 13–20. <https://doi.org/10.17509/seict.v6i1.86210>
- Khatib Sulaiman, J., Reni Dwi Astuti, A., Cahyono, N., & Amikom Yogyakarta, U. (2023). Analisis Topic Modelling Persepsi Pengguna Internet Menggunakan Metode Latent Dirichlet Allocation. *Indonesian Journal of Computer Science Attribution*, 12(1), 2023–2326.
- Khofifah, W., Rahayu, D. N., & Yusuf, A. M. (2022). Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 16(4), 28–38. <https://doi.org/10.35969/interkom.v16i4.192>
- Kim, K., Han, S. L., Jang, Y. Y., & Shin, Y. C. (2020). The effects of the antecedents of “buy-online-pick-up-in-store” service on consumer’s bopis choice behaviour. *Sustainability (Switzerland)*, 12(23), 1–15. <https://doi.org/10.3390/su12239989>
- Lasarudin, A., & Yoanda, P. P. (2024). Implementasi importance performance analysis (IPA) untuk mengetahui kualitas pelayanan terhadap harapan wisatawan Al-Fatih Tour & Travel. *Jurnal Pariwisata Pesona*, 9(1), 107–115. <https://doi.org/10.26905/jpp.v9i1.13188>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Mehenaoui, Z., Merabti, C., Tadjer, H., & Lafifi, Y. (2024). *A Comparative Study On Sentiment Lexicons For Automatic Labeling*.

- Murdy, S. (2012). Perceptions of visitor relationship marketing opportunities by destination marketers: An Importance-Performance Analysis. *Tourism Management*, 33, 1281–1285.
- Nurlayli, A., & Nasichuddin, M. A. (2019). Topic Modeling Penelitian Dosen JPTEI UNY pada Google Scholar Menggunakan Latent Dirichlet Allocation. *ELINVO(Electronics, Informatics, and Vocational Education)*, 4(2), 154–161. <https://doi.org/10.21831/elinvo.v4i2>
- Pau, A. I. K. (2025, November 2). *Indonesia Masuk 4 Besar Pengguna Instagram Terbanyak Dunia*. Radio Republik Indonesia. <https://rri.co.id/kupang/hiburan/1946081/indonesia-masuk-4-besar-pengguna-instagram-terbanyak-dunia>
- Pérez, F., Granger, B. E., & Hunter, J. D. (2011). Python An Ecosystem for Scientific Compu. *IEEE Xplore*, 13(2).
- Pratiwi, I. W. (2018). IMPLEMENTASI IMPORTANCE-PERFORMANCE ANALYSIS (IPA): ANALISIS PREFERENSI KONSUMEN MUDA MIE SETAN MALANG. *Jurnal Imiah Mahasiswa FEB*, 6.
- Purnama, D. A., Lumi, D. P., Febriani, A., Utama T, A. R., Salindri, S. D., Anugerah, A. R., & Aliafari, N. (2025). Data-driven public policy for electric vehicles (EV) through open innovation and dynamic consumer preferences: A time-series social media analysis using integrated IPA-product improvability model. *Journal of Open Innovation: Technology, Market, and Complexity*, 11(3). <https://doi.org/10.1016/j.joitmc.2025.100583>
- Purnama, D. A., Subagyo, & Masruroh, N. A. (2023). Online data-driven concurrent product-process-supply chain design in the early stage of new product development. *Journal of Open Innovation: Technology, Market, and Complexity*, 9(3). <https://doi.org/10.1016/j.joitmc.2023.100093>
- Rahman, A., Sofiyah, E. S., Sianipar, I. M. J., Edwisafira, A., Suhardono, S., Lee, C. H., Nguyen, V. V., & Suryawan, I. W. K. (2025). Evaluating tourist importance-performance and support for development in Komodo National Park. *Environmental Challenges*, 20. <https://doi.org/10.1016/j.envc.2025.101229>
- Rudini, D., Gita Purnama, D., & Achmad Khan, A. (2023). PENGGUNAAN TEKNIK WEB SCRAPING DALAM APLIKASI PENGAMBILAN DATA DARI GOOGLE MAPS UNTUK MENUNJANG DIGITAL MARKETING. *Lentera: Multidisciplinary Studies*, 2(1). <https://lentera.publikasiku.id/index.php>
- Satrio Wibowo, M., & Arviana Belia, L. (2023). Partisipasi Masyarakat dalam Pengembangan Pariwisata Berkelanjutan. *Jurnal Manajemen Perhotelan Dan Pariwisata*, 6(1), 2023.
- Siregar, A. A. (2025). An Analytical Approach to Visitor Carrying Capacity and Conservation: The Pindul Cave of Yogyakarta. *International Journal of Green Tourism Research and Applications*, 7(1), 49–59. <https://doi.org/10.31940/ijogtra.v7i1.49-59>
- Sudiatmika, I. P. G. A., Satya Saputra, P., Rahardian, R., & Santhi Dewi, K. (2025). *Sentiment analysis of tourist reviews on google maps for pura besakih using machine learning algorithms*. 14(1), 149–158. www.ejournal.isha.or.id/index.php/Mandiri
- Sukmaraga, L. (2016). *Arahan Pengembangan Produk Wisata Kawasan Goa Pindul Berdasarkan Pendapat Pengunjung*. Universitas Gadjah Mada.
- Suprihatin, E., & Rofiq, A. (2023). ANALISIS KUALITAS PELAYANAN DAN PENTINGNYA KINERJA TERHADAP PENCAPAIAN KEBERLANJUTAN PENGEMBANGAN PARIWISATA KINERJA LAUT PARIWISATA. *Pendidikan Dan Humaniora*, 7(2). <https://doi.org/10.36526/js.v3i2.3209>
- Wirawan, P. E., & Semara, I. M. T. (2021). *Pengantar Pariwisata* (A. A. A. S. Arianty, Ed.). INSTITUT PARIWISATA DAN BISNIS INTERNASIONAL. www.stpbi.ac.id

Yuniar, E., Utsalinah, D. S., & Wahyuningsih, D. (2022). Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompok Digital dengan Menggunakan Algoritma Machine Learning. *Jurnal Janitra Informatika Dan Sistem Informasi*, 2(1), 35–42.
<https://doi.org/10.25008/janitra.v2i1.145>

LAMPIRAN

A-Pengolahan Data *Google Maps* dan *Instagram*

```

[1] from apify_client import ApifyClient
Python

[2] import pandas as pd
Python

# Initialize the ApifyClient with your API token
client = ApifyClient("apify_api_")

# Prepare the Actor input
run_input = {
    "postUrls": [
        "https://www.instagram.com/reel/DEjz8lOT3gY/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==",
        "https://www.instagram.com/reel/DEeFCDHzh-1/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==",
        "https://www.instagram.com/p/DDTCuxHzlZ7/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==",
    ],
    "maxCommentsPerPost": 200,
    "sortOrder": "popular",
}

# Run the Actor and wait for it to finish
run = client.actor("499mNnuVGkU2S5rh1").call(run_input=run_input)

# Fetch and print Actor results from the run's dataset (if there are any)
items = []
for item in client.dataset(run["defaultDatasetId"]).iterate_items():
    items.append(item)

df = pd.DataFrame(items)
df

df.to_excel('ScrapIG3.xlsx', index=False)
Python

[1] !pip install pandas
!pip install gensim
!pip install spacy
!pip install matplotlib
!pip install nltk
!pip install openpyxl
> ... Show hidden output

[1] from google.colab import files
uploaded = files.upload()
> Show hidden output

[1] import nltk

nltk.download('punkt')
nltk.download('punkt_tab')

import pandas as pd
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer
> ... Show hidden output

```

```
[ ] # baca excel
df = pd.read_excel("04_INI DIA.xlsx", sheet_name="GMAPS")

# ambil kolom
texts = df["Text Stopword"].dropna()

[ ] # 1. TOKENISASI
# -----
df["Tokens"] = df["Text Stopword"].apply(
    lambda x: word_tokenize(x) if isinstance(x, str) else []
)

[ ] # 2. STEMMING
# -----
from nltk.stem import PorterStemmer
stemmer = PorterStemmer()

df["Stemming"] = df["Tokens"].apply(
    lambda tokens: [stemmer.stem(w) for w in tokens] if isinstance(tokens, list) else []
)

import gensim
import matplotlib.pyplot as plt
from gensim.corpora import Dictionary
from gensim.models import LdaModel
from gensim.models import CoherenceModel

import pandas as pd
import numpy as np
import re
import ast

documents = df["Stemming"].tolist()

#Buat kamus kata
dictionary = Dictionary(documents)
corpus = [dictionary.doc2bow(doc) for doc in documents]

k_range = range(10, 21)
coherence_scores = []

for k in k_range:
    lda_temp = LdaModel(
        corpus=corpus,
        id2word=dictionary,
        num_topics=k,
        passes=10,
        random_state=100
    )

    coherence_model = CoherenceModel(
        model=lda_temp,
        texts=documents,
        dictionary=dictionary,
        coherence='c_v'
    )

    coherence_scores.append(coherence_model.get_coherence())

#tentukan nilai k dengan nilai Koherensi tertinggi
best_k = k_range[coherence_scores.index(max(coherence_scores))]

plt.plot(list(k_range), coherence_scores)
plt.xticks(k_range)
plt.scatter(best_k, max(coherence_scores), color='red', label='Best k = {}'.format(best_k)) #Tambahkan marker untuk nilai k terbaik
plt.xlabel("Jumlah Topik")
plt.ylabel("Nilai Koherensi")
plt.title("Grafik Koherensi vs Jumlah Topik")
plt.show()

Show hidden output

#tampilkan nilai koherensi dari best_k
best_k_index = best_k - 10 #set indek mulai dari 2 karena k range mulai dari 2
best_coherence_score = coherence_scores[best_k_index]
print("Nilai Koherensi untuk best_k (k = {}) : {}".format(best_k, best_coherence_score))

Show hidden output

lda_model = LdaModel(
    corpus=corpus,
    id2word=dictionary,
    num_topics=best_k,
    passes=10,
    random_state=100
)
```

```

for i in range(best_k):
    print(f"\ntopik {i}:")
    print(lda_model.print_topic(i, topn=10))

doc_topics = [lda_model.get_document_topics(doc) for doc in corpus]

dominant_topics = [
    max(topics, key=lambda x: x[1])[0] if topics else None
    for topics in doc_topics
]

df['Topik'] = dominant_topics

print("Topik unik:", sorted(df['Topik'].dropna().unique()))

topic_counts = df['Topik'].value_counts().sort_index()

importance = topic_counts / topic_counts.sum()

importance_df = importance.reset_index()
importance_df.columns = ['Topik', 'Importance']

print(importance_df)

```

```

import nltk
nltk.download('vader_lexicon')

from nltk.sentiment.vader import SentimentIntensityAnalyzer

[nltk_data] Downloading package vader_lexicon to /root/nltk_data...

sia = SentimentIntensityAnalyzer()

import re

def clean_text(text):
    text = str(text).lower()          # lowercase
    text = re.sub(r"http[s]+", "", text) # hapus link
    text = re.sub(r"@w+", "", text)    # hapus mention
    text = re.sub(r"#w+", "", text)    # hapus hashtag
    return text

df['clean_text'] = df['Text Translate'].apply(clean_text)

df['sentiment_score'] = df['clean_text'].apply(lambda x: sia.polarity_scores(x)['compound'])

df[['Text Translate', 'sentiment_score']].head()

```

... Show hidden output

```

def label_sentiment(score):
    if score >= 0.05:
        return 'Positif'
    elif score <= -0.05:
        return 'Negatif'
    else:
        return 'Netral'

df['sentiment_label'] = df['sentiment_score'].apply(label_sentiment)

df['sentiment_label'].value_counts()

Show hidden output

satisfaction = df.groupby('Topik')['sentiment_score'].mean().reset_index()
satisfaction.columns = ['Topik', 'Satisfaction']

print(satisfaction)

Show hidden output

ipa_df = importance_df.merge(satisfaction, on='Topik')
print(ipa_df)

Show hidden output

```

```

1 df[['Topik', 'sentiment_score']].head()
> Show hidden output
1 df.to_excel("INI DIA GMAPSI.xlsx", index=False)
from google.colab import files
files.download("INI DIA GMAPSI.xlsx")
> ... Show hidden output

```