

**ANALISIS SENTIMEN DAN PEMODELAN TOPIK ULASAN
WISATAWAN MENGGUNAKAN *INSET LEXICON* DAN LDA
UNTUK MENDUKUNG PENENTUAN PRIORITAS PERBAIKAN
LAYANAN PADA HEHA *FOREST***

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Program Studi Teknik Industri - Fakultas Teknologi Industri
Universitas Islam Indonesia**



Nama : Khalid Farrossyah Asri Putra
No. Mahasiswa : 22522055

**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, 30 Mei 2026



A handwritten signature in black ink, appearing to read 'Kus', written over the banknote.

(Khalid Farrossyah Asri Putra)

NIM: 22522055

SURAT BUKTI PENELITIAN



FAKULTAS
TEKNOLOGI INDUSTRI

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

SURAT KETERANGAN PENELITIAN

Nomor: 019/Ka.Lab.Datmin/70/Lab.Datmin/VI/2026

Assalamu'alaikum Warahmatullahi Wabarakatuh

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

Nama : Khalid Farossyah Asri Putra
No. Mhs : 22522055
Dosen Pembimbing : Dr. Dwi Adi Purnama, S.T.

Telah selesai melaksanakan penelitian yang berjudul "ANALISIS SENTIMEN DAN PEMODELAN TOPIK ULASAN WISATAWAN MENGGUNAKAN INSET LEXICON DAN LDA UNTUK MENDUKUNG PENENTUAN PRIORITAS PERBAIKAN LAYANAN PADA HEHA FOREST" di Laboratorium Data Mining, Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 16 Maret sampai dengan tanggal 30 April 2026.

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

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Yogyakarta, 02 Juni 2026

Kepala Laboratorium
Data Mining

Dr. Dwi Adi Purnama, S.T.

LEMBAR PENGESAHAN PEMBIMBING
ANALISIS SENTIMEN DAN PEMODELAN TOPIK ULASAN
WISATAWAN MENGGUNAKAN *INSET LEXICON* DAN LDA
UNTUK MENDUKUNG PENENTUAN PRIORITAS PERBAIKAN
LAYANAN PADA HEHA *FOREST*



Dosen Pembimbing

(Dr. Dwi Adi Purnama, S.T.)

LEMBAR PENGESAHAN DOSEN PENGUJI
ANALISIS SENTIMEN DAN PEMODELAN TOPIK ULASAN
WISATAWAN MENGGUNAKAN *INSET LEXICON* DAN LDA
UNTUK MENDUKUNG PENENTUAN PRIORITAS PERBAIKAN
LAYANAN PADA HEHA *FOREST*

TUGAS AKHIR

Disusun Oleh :

Nama : Khalid Farrossyah Asri Putra

No. Mahasiswa : 22 522 055

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, 10 - Juli – 2026

Tim Penguji

Dr. Dwi Adi Purnama, S.T.

Ketua

Elanjati Worldailmi, S.T., M.Sc.

Anggota I

Sayyidah Maulidatul Afraah, S.T., M.T.

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

Bismillahirrahmanirrahim, atas izin Allah SWT saya mempersembahkan Tugas Akhir ini untuk Ayah dan Ibu saya, Ayah Asri Yonalsyah, A.P., M.Si., dan Ibu Febriyenti S.E., M.M., untuk segala doa, cinta, dan perjuangannya untuk memberikan yang terbaik karena hal itu yang bisa membuat saya bisa sampai ke titik ini.

MOTTO

"Karena sesungguhnya sesudah kesulitan itu ada kemudahan."

(QS.Al-Insyirah: 5-6)

"Ridha Allah tergantung pada ridha orang tua dan murka Allah tergantung pada murka orang tua."

(HR. At-Tirmidzi, Ibnu Hibban, dan Al-Hakim)

"Lakamatī lā yumkinu ikhfā'uhā."

(Aldous)

KATA PENGANTAR

Bismillahirrahmanirrahim

Assalamualaikum Warahmatullahi Wabarakatuh

Alhamdulillah segala puji dan syukur penulis panjatkan ke hadirat Allah SWT yang senantiasa melimpahkan rahmat, karunia, dan hidayah-Nya. Berkat rahmat tersebut, penulis dapat menyelesaikan Tugas Akhir. Tidak lupa, shalawat serta salam penulis haturkan kepada Nabi Muhammad SAW, suri teladan seluruh umat manusia, yang telah memberikan petunjuk dan ajaran yang menjadi panduan hidup bagi umat Islam. Semoga segala upaya yang telah dilakukan dalam proses ini mendapat ridho-Nya dan dapat memberikan manfaat bagi penulis dan pihak-pihak yang terkait. Terima kasih juga penulis sampaikan kepada semua pihak yang telah memberikan dukungan selama pelaksanaan Tugas Akhir ini.

Selama penyusunan Tugas Akhir, penulis mendapatkan banyak bimbingan, arahan, dukungan, serta bantuan dari berbagai pihak yang berkontribusi, baik secara langsung maupun tidak langsung. Atas segala dukungan tersebut, penulis dengan tulus mengucapkan terima kasih kepada semua pihak yang telah memberikan bantuan, arahan, dan dorongan hingga tersusunnya Tugas Akhir ini, antara lain:

1. Ibu Prof. Dr. Sri Kusumadewi, S.Si., M.T., selaku Dekan Fakultas Teknologi Industri Universitas Islam Indonesia, yang telah memberikan arahan kebijakan akademik selama program ini berlangsung.
2. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM., selaku Ketua Program Studi Teknik Industri Program Sarjana Fakultas Teknologi Industri Universitas Islam Indonesia, atas dukungan dan kebijakan yang memfasilitasi pelaksanaan Tugas Akhir ini.
3. Bapak Dr. Dwi Adi Purnama S.T., selaku dosen pembimbing Tugas Akhir yang dengan penuh kesabaran dan perhatian memberikan bimbingan, kritik, dan saran yang sangat membantu dalam penyelesaian Tugas Akhir ini.
4. Kedua orang tua tercinta, Ayah Asri Yonalsyah, A.P., M.Si., dan Ibu Febriyenti S.E., M.E., penulis yang selalu memberikan doa, dukungan, perhatian, serta semangat yang tak pernah surut.
5. Seluruh sahabat pada grup Apenih dan Pamungkas, Nasywa, Abi, Allan, Ceri, Cebong, Darsul, Duta, Gean, Iyan, Nandes, NW, Podol, Raka, Regung, Ryan, dan Tengkot yang telah menemani penulis selama masa kuliah.
6. Seluruh sahabat di grup Netral Brothers dan 22 *Full* Petarung yang telah mendukung penulis selama masa kuliah.
7. Untuk semua pihak yang tidak bisa disebutkan satu persatu yang telah memberikan bantuan dan kontribusi selama masa perkuliahan.

Penulis menyadari sepenuhnya bahwa Tugas Akhir ini masih jauh dari kata sempurna, mengingat keterbatasan pengetahuan dan pengalaman yang dimiliki. Penulis juga menyadari bahwa masih terdapat banyak kekurangan dalam laporan ini. Oleh karena itu, segala bentuk kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang. Harapannya, laporan ini dapat memberikan manfaat sesuai dengan tujuannya dan menjadi sumber tambahan wawasan serta pengetahuan bagi para pembaca. Penulis juga berharap, semua kebaikan, bimbingan, dan bantuan yang telah diberikan oleh berbagai pihak akan mendapatkan balasan yang setimpal dari Allah SWT. Semoga

laporan ini tidak hanya menjadi dokumentasi semata, tetapi juga dapat memberikan nilai tambah yang bermanfaat bagi dunia akademik dan profesional.

Wassalamualaikum Warahmatullahi Wabarakatuh

Sleman, 20 Mei 2026



Khalid Farrossyah Asri Putra

NIM 22522055

ABSTRAK

Perkembangan sektor pariwisata di Indonesia, khususnya di Kabupaten Sleman, Daerah Istimewa Yogyakarta, menunjukkan peningkatan jumlah wisatawan yang signifikan dari tahun ke tahun. Salah satu destinasi wisata yang mengalami perkembangan pesat adalah HeHa Forest, yang memadukan konsep wisata alam, restoran, *Spot* foto *Instagramable*, *playground*, dan *mini zoo*. Tingginya aktivitas wisatawan di media digital seperti *Google Maps* dan *Instagram* menghasilkan berbagai ulasan yang mencerminkan persepsi serta pengalaman pengunjung. Penelitian ini bertujuan untuk mengidentifikasi topik utama ulasan wisatawan, menganalisis kecenderungan sentimen pengunjung, menentukan prioritas aspek layanan yang perlu diperbaiki, serta memberikan rekomendasi strategi pengembangan wisata pada HeHa Forest. Metode yang digunakan dalam penelitian ini meliputi *web scraping* untuk pengumpulan data ulasan dari *Google Maps* dan *Instagram* periode 2023 hingga Februari 2026, *text mining* sebagai tahap *pre-processing* data, analisis sentimen menggunakan *InSet Lexicon*, *topic modelling* menggunakan *Latent Dirichlet Allocation* (LDA), serta metode *Technique for Order Preference by Similarity to Ideal Solution* (TOPSIS) untuk menentukan prioritas perbaikan layanan. Hasil penelitian menunjukkan bahwa topik utama yang paling sering dibahas wisatawan meliputi fasilitas, harga, pelayanan, makanan, *Spot* foto, dan pengalaman wisata secara keseluruhan. Analisis sentimen menunjukkan bahwa pada *Google Maps* topik 4 memiliki nilai sentimen tertinggi sebesar 1,302, sedangkan pada *Instagram* topik 1 memiliki nilai sentimen tertinggi sebesar 0,07059. Hasil perhitungan TOPSIS menunjukkan bahwa pada *Google Maps*, topik 3 yang merupakan *Spot* Foto Menarik terpilih sebagai prioritas utama perbaikan dengan nilai preferensi sebesar 0,77952, sedangkan pada *Instagram* topik 7 yang merupakan Biaya Aktivitas Keluarga menjadi prioritas utama dengan nilai preferensi sebesar 0,81089. Penelitian ini menghasilkan rekomendasi strategi pengembangan layanan yang dapat digunakan pengelola HeHa Forest untuk meningkatkan kualitas layanan, kepuasan pengunjung, serta daya saing destinasi wisata secara berkelanjutan.

Kata Kunci: Analisis Sentimen, *InSet Lexicon*, *Latent Dirichlet Allocation*, TOPSIS, HeHa Forest .

DAFTAR ISI

PERNYATAAN KEASLIAN.....	i
SURAT BUKTI PENELITIAN	ii
LEMBAR PENGESAHAN PEMBIMBING	iii
LEMBAR PENGESAHAN DOSEN PENGUJI.....	iv
HALAMAN PERSEMBAHAN	v
MOTTO.....	vi
KATA PENGANTAR.....	vii
ABSTRAK.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	7
1.3 Tujuan Penelitian	7
1.4 Manfaat Penelitian	7
1.5 Batasan Penelitian.....	8
BAB II TINJAUAN PUSTAKA	9
2.1 Kajian Literatur.....	9
2.1.1 Web Scrapping dalam Data Pariwisata	9
2.1.2 Analisis Sentimen dalam Penelitian Pariwisata	9
2.1.3 Topic Modelling Menggunakan Latent Dirichlet Allocation (LDA)....	10
2.1.4 Metode Pengambilan Keputusan TOPSIS	10
2.1.5 Integrasi Metode dalam Penelitian Wisata.....	10
2.1.6 Research Gap dan Posisi Penelitian	11
2.2 Landasan Teori	12
2.2.1 Pariwisata	12
2.2.2 Web Scraping	13
2.2.3 Text Mining.....	13
2.2.4 User Generated Content	14
2.2.6 InSet Lexicon	14
2.2.5 Analisis Sentimen.....	16
2.2.7 Topic Modelling.....	16
2.2.8 Latent Dirichlet Allocation (LDA).....	17
2.2.9 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)	18
2.2.10 Integrasi Metode.....	20
BAB III METODE PENELITIAN	21
3.1 Subjek Penelitian	21
3.2 Objek Penelitian.....	21
3.3 Jenis & Sumber Data	21
3.4 Alur Penelitian	21
BAB IV PENGUMPULAN DAN PENGOLAHAN DATA.....	27
4.1 Pengumpulan Data.....	27
4.1.1 Data Google Maps.....	27
4.1.2 Data Instagram	29
4.2 <i>Pre-processing</i> Data	36
4.2.1 Translasi	36

4.2.2	Stopword	36
4.2.3	Pengubahan huruf kecil (Case Folding)	36
4.2.4	Tokenisasi	36
4.2.5	Penghapusan data subjektif, karakter non-alphabet (filtering)	37
4.2.6	Stemming	37
4.2.7	Penyusunan kalimat	37
4.3	Pemodelan Topik (<i>Latent Dirichlet Allocation/LDA</i>)	37
4.3.1	Hasil pemodelan topik	38
4.4	Interpretasi Topik	43
4.4.1	Interpretasi topik pada data Google Maps	43
4.4.2	Interpretasi topik pada data Instagram	45
4.5	Analisis Sentimen menggunakan <i>InSet Lexicon</i>	47
4.5.1	Metode analisis sentiment	47
4.5.2	Hasil analisis sentiment Instagram	47
4.5.3	Hasil analisis sentiment Google Maps	48
4.6	TOPSIS	48
4.6.1	Penentuan alternatif dan kriteria	48
4.6.2	Pembentukan matriks Keputusan	49
4.6.3	Normalisasi dan pembobotan	50
4.6.4	Penentuan solusi ideal dan nilai prefensi	53
BAB V PEMBAHASAN		59
5.1	Perbandingan Hasil Analisis antar <i>Platform</i>	59
5.2	Hasil Prioritas TOPSIS	60
5.3	Rekomendasi Perbaikan Menggunakan Pendekatan <i>Lotus Blossom</i>	62
5.3.1	Rekomendasi perbaikan Instagram	63
5.3.2	Rekomendasi perbaikan Google Maps	64
5.4	Visualisasi Konsep Implementasi Rekomendasi	64
5.4.1	Visualisasi konsep Instagram	65
5.4.2	Visualisasi konsep Google Maps	69
BAB VI PENUTUP		74
6.1	Kesimpulan	74
6.2	Saran	75
DAFTAR PUSTAKA		76

DAFTAR TABEL

Tabel 1. 1 SWOT	4
Tabel 2. 1 Posisi Penelitian.....	12
Tabel 2. 2 Tabel perbandingan metode analisis sentimen	15
Tabel 2. 3 Tabel perbandingan <i>topic modelling</i>	17
Tabel 4. 1 Data <i>Review Google Maps</i>	28
Tabel 4. 2 Data <i>comment instagram</i>	30
Tabel 4. 3 Translasi.....	36
Tabel 4. 4 Data <i>stopword</i>	36
Tabel 4. 5 Data <i>case folding</i>	36
Tabel 4. 6 Data tokenisasi.....	37
Tabel 4. 7 Data <i>Filtering</i>	37
Tabel 4. 8 Data <i>Stemmed</i>	37
Tabel 4. 9 Data Penyusunan Kalimat.....	37
Tabel 4. 10 Topik <i>Google Maps</i>	40
Tabel 4. 11 <i>Topic 10 Instagram</i>	41
Tabel 4. 12 Interpretasi topik pada data <i>Google Maps</i>	43
Tabel 4. 13 Interpretasi topik pada <i>Instagram</i>	45
Tabel 4. 14 Rekapitulasi sentimen <i>Instagram</i>	47
Tabel 4. 15 Rekapitulasi sentimen <i>Google Maps</i>	48
Tabel 4. 16 Matriks Keputusan <i>Google Maps</i>	49
Tabel 4. 17 Matriks Keputusan <i>Instagram</i>	50
Tabel 4. 18 Matriks normalisasi <i>Google Maps</i>	51
Tabel 4. 19 Matriks ternormalisasi terbobot <i>Google Maps</i>	51
Tabel 4. 20 Matriks normalisasi <i>Instagram</i>	52
Tabel 4. 21 Matriks ternormalisasi terbobot <i>Instagram</i>	53
Tabel 4. 22 Hasil jarak <i>Google Maps</i>	55
Tabel 4. 23 Hasil jarak <i>Instagram</i>	56
Tabel 4. 24 <i>Rank</i> nilai prefensi <i>Google Maps</i>	57
Tabel 4. 25 <i>Rank</i> nilai prefensi <i>Instagram</i>	58

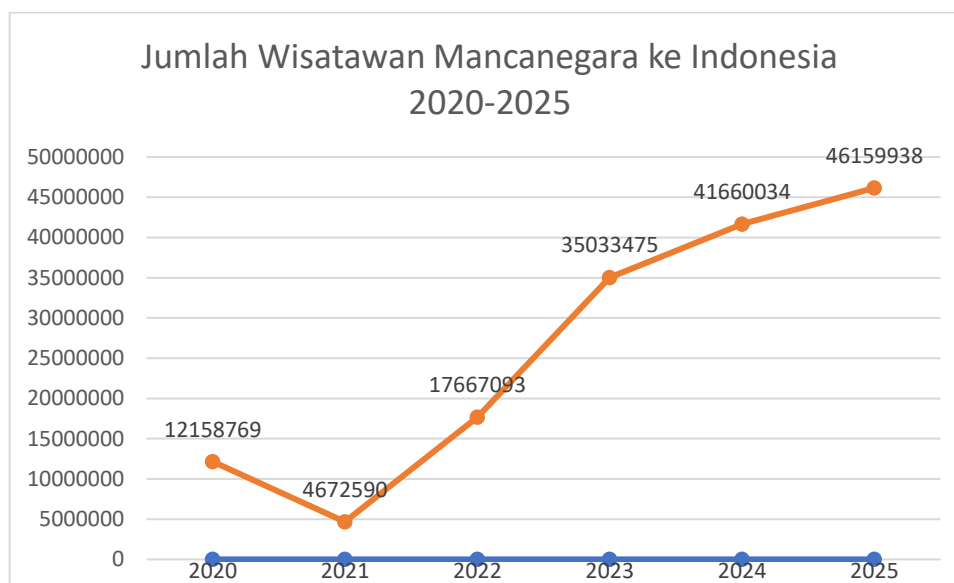
DAFTAR GAMBAR

Gambar 1. 1 Jumlah Wisatawan Mancanegara ke Indonesia Sumber : (Badan Pusat Statistik, 2026).....	1
Gambar 1. 2 Jumlah Wisatawan Nusantara D. I. Yogyakarta Sumber : (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026).....	2
Gambar 1. 3 Jumlah Wisatawan Nusantara D. I. Yogyakarta pada Kabupaten atau Kota 2020 - 2025 Sumber: (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026).....	3
Gambar 1. 4 Jumlah Wisatawan Nusantara D. I. Yogyakarta pada Kabupaten atau Kota 2020 - 2025 Sumber : (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026).....	3
Gambar 3. 1 Alur penelitian	22
Gambar 4. 1 Grafik <i>coherence Google Maps</i>	38
Gambar 4. 2 Grafik <i>coherence Instagram</i>	39
Gambar 5. 1 <i>Lotus Blossom Instagram</i>	63
Gambar 5. 2 <i>Lotus Blossom Google Maps</i>	64
Gambar 5. 3 Konsep visual Informasi harga makanan dan aktivitas di Media Sosial ...	66
Gambar 5. 4 Konsep visual area bermain gratis untuk anak	67
Gambar 5. 5 Paket keluarga hemat	68
Gambar 5. 6 <i>Spot</i> foto interaktif	70
Gambar 5. 7 Perawatan area foto.....	71
Gambar 5. 8 Pencahayaan area foto.....	72

BAB I PENDAHULUAN

1.1 Latar Belakang

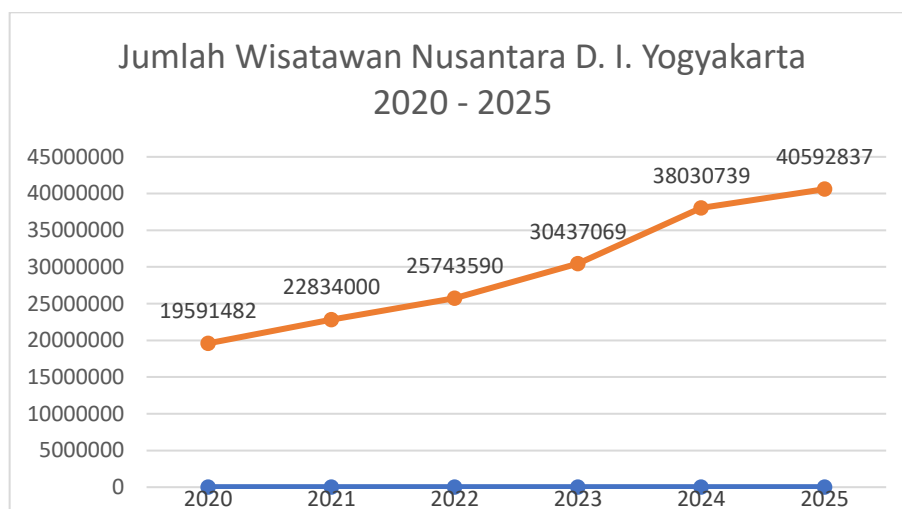
Sektor pariwisata telah berkembang pesat selama beberapa tahun terakhir dan terbukti menjadi salah satu industri paling populer di dunia. Pendapatan sektor ini diproyeksikan mencapai US\$854,80 miliar pada tahun 2023, dengan tingkat pertumbuhan tahunan sebesar 4,41% dari 2023 hingga 2027, dan menciptakan *volume* pasar sebesar US\$1.016,00 miliar pada tahun 2027 (Hasibuan dkk., 2023). Menurut data resmi Badan Pusat Statistik pada tahun 2020 hingga 2025, jumlah wisatawan mancanegara ke Indonesia menunjukkan peningkatan yang signifikan dari 12,1 juta pada tahun 2020 hingga 46,1 juta pada tahun 2025. Kenaikan tersebut menunjukkan bahwa meningkatnya sektor pariwisata di Indonesia beberapa tahun terakhir walau terdapat penurunan pada tahun 2021 namun berhasil naik Kembali di tahun 2022. Tren peningkatan jumlah wisatawan mancanegara ke Indonesia dapat dilihat pada Gambar 1. 1.



Gambar 1. 1 Jumlah Wisatawan Mancanegara ke Indonesia
Sumber : (Badan Pusat Statistik, 2026)

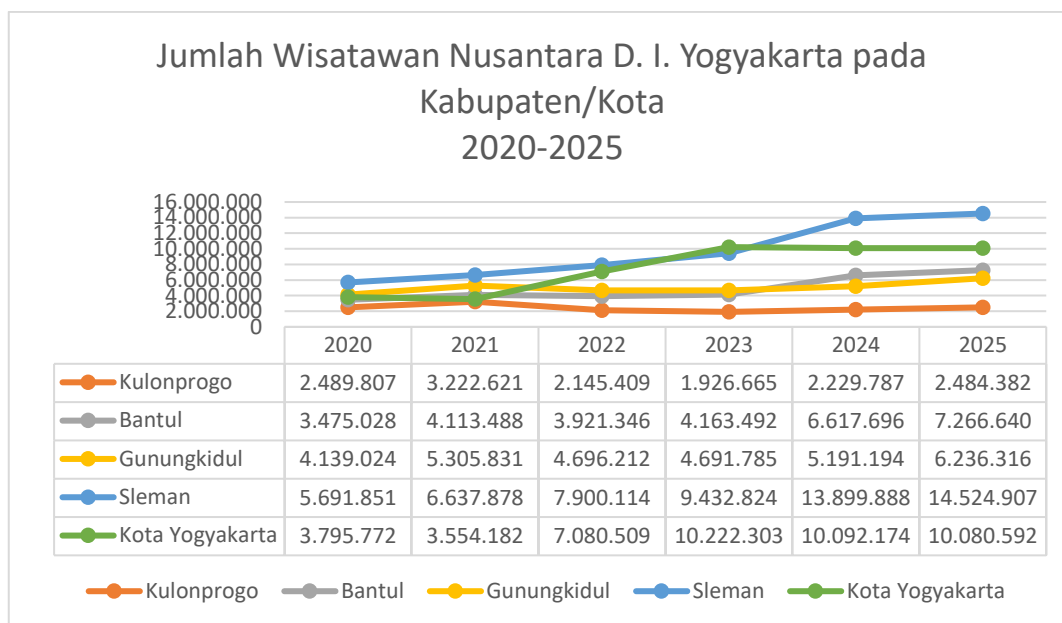
Berdasarkan Gambar 1. 1, jumlah wisatawan mancanegara yang berkunjung ke Indonesia menunjukkan tren peningkatan dari tahun ke tahun. Kondisi tersebut mengindikasikan bahwa daya tarik destinasi wisata di Indonesia terus mengalami peningkatan. Salah satu daerah tujuan wisata yang turut mengalami peningkatan kunjungan adalah Daerah Istimewa Yogyakarta yang dikenal sebagai salah satu destinasi

wisata unggulan di Indonesia. Perkembangan jumlah wisatawan di Provinsi Daerah Istimewa Yogyakarta dapat dilihat pada Gambar 1. 2.



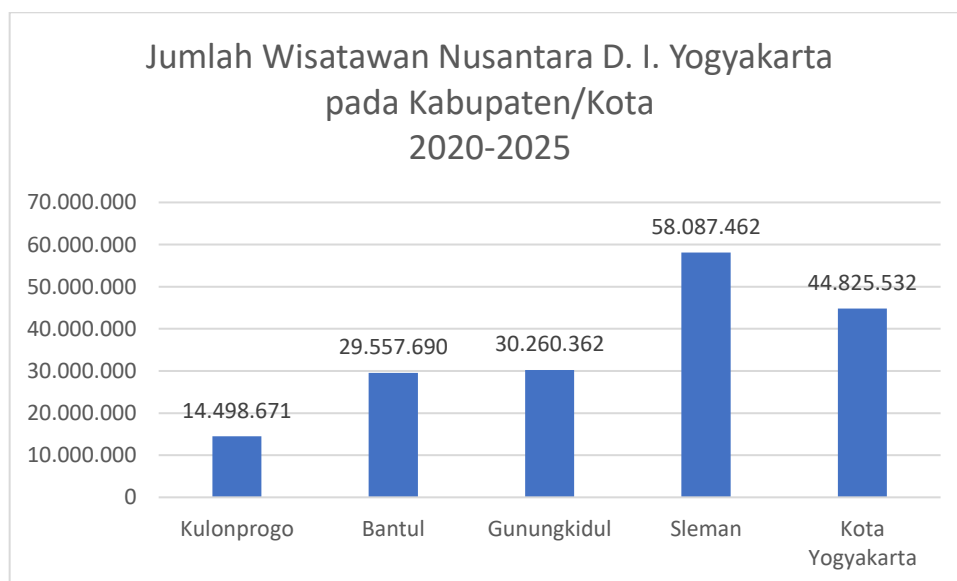
Gambar 1. 2 Jumlah Wisatawan Nusantara D. I. Yogyakarta
Sumber : (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026)

Berdasarkan Gambar 1. 2, jumlah wisatawan nusantara di Provinsi Daerah Istimewa Yogyakarta mengalami peningkatan dari 19,5 juta kunjungan pada tahun 2020 menjadi 40,5 juta kunjungan pada tahun 2025. Peningkatan tersebut menunjukkan bahwa minat masyarakat untuk berkunjung ke berbagai destinasi wisata di Daerah Istimewa Yogyakarta terus meningkat. Namun, peningkatan tersebut tidak terjadi secara merata pada seluruh kabupaten dan kota. Oleh karena itu, diperlukan identifikasi distribusi kunjungan wisatawan pada masing-masing wilayah yang disajikan pada Gambar 1. 3.



Gambar 1. 3 Jumlah Wisatawan Nusantara D. I. Yogyakarta pada Kabupaten atau Kota 2020 - 2025

Sumber: (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026)



Gambar 1. 4 Jumlah Wisatawan Nusantara D. I. Yogyakarta pada Kabupaten atau Kota 2020 - 2025

Sumber : (Badan Pusat Statistik Provinsi DI Yogyakarta, 2026)

Data pada Gambar 1.3 dan Gambar 1.4 menunjukkan bahwa jumlah wisatawan yang berkunjung ke Kabupaten Sleman meningkat dari tahun 2020 hingga 2025 sehingga Kabupaten Sleman menjadi kabupaten dengan jumlah kunjungan wisatawan tertinggi, yaitu mencapai 58 juta kunjungan. Kondisi tersebut menunjukkan bahwa Kabupaten Sleman memiliki daya tarik wisata yang lebih tinggi dibandingkan kabupaten dan kota

lain di Provinsi Daerah Istimewa Yogyakarta. Banyaknya jumlah kunjungan tersebut juga mendorong munculnya berbagai destinasi wisata baru maupun pengembangan destinasi yang telah ada. Salah satu destinasi yang berkembang dan memiliki popularitas tinggi adalah HeHa *Forest*.

HeHa *Forest* dipilih sebagai objek penelitian karena merupakan salah satu destinasi wisata di Kabupaten Sleman yang memiliki popularitas tinggi di media sosial dan *platform* digital, seperti *Google Maps* dan *Instagram*. Destinasi ini tidak hanya menawarkan wisata alam, tetapi juga menggabungkan konsep restoran, *spot* foto *instagramable*, *playground*, dan *mini zoo* sehingga menghasilkan beragam opini dari pengunjung. Keberagaman opini tersebut menunjukkan bahwa HeHa *Forest* memiliki berbagai faktor yang memengaruhi pengalaman pengunjung, baik berupa kelebihan maupun kekurangan. Oleh karena itu, pendekatan SWOT digunakan sebagai gambaran awal untuk mengidentifikasi kondisi internal dan eksternal HeHa *Forest*.

Tabel 1. 1 SWOT

Aspek	Penjelasan
Strengths	Memiliki konsep wisata alam yang dipadukan dengan restoran, <i>Spot</i> foto <i>Instagramable</i> , <i>playground</i> , dan <i>mini zoo</i> sehingga menarik minat wisatawan, khususnya generasi muda dan pengguna media sosial.
Weaknesses	Terdapat <i>gap</i> antara ekspektasi dan pengalaman pengunjung yang memunculkan berbagai opini negatif terhadap layanan, fasilitas, harga, maupun pengalaman wisata.
Opportunities	Tingginya pertumbuhan sektor pariwisata di Yogyakarta dan Kabupaten Sleman membuka peluang peningkatan jumlah wisatawan serta pengembangan kualitas layanan berbasis opini digital pengunjung.
Threats	Penyebaran opini negatif di media sosial dan <i>platform</i> ulasan digital dapat mempengaruhi citra HeHa <i>Forest</i> serta keputusan calon wisatawan untuk berkunjung.

Berdasarkan analisis pada Tabel 1.1, dapat dilihat bahwa HeHa *Forest* memiliki daya tarik wisata yang kuat, namun masih terdapat berbagai kritik dan keluhan yang disampaikan pengunjung melalui media digital. Hasil analisis SWOT tersebut memberikan gambaran umum mengenai kondisi objek wisata, tetapi belum mampu menunjukkan secara rinci permasalahan yang paling banyak dibahas maupun aspek yang

perlu diprioritaskan untuk diperbaiki. Oleh karena itu, diperlukan analisis lebih lanjut terhadap opini pengunjung menggunakan metode LDA, analisis sentimen, dan TOPSIS untuk mengidentifikasi topik pembahasan, mengetahui kecenderungan sentimen, serta menentukan prioritas perbaikan layanan.

HeHa *Forest* merupakan salah satu destinasi wisata baru yang hadir setelah pengembangan HeHa *Sky View* dan HeHa *Ocean View*. Destinasi ini mengusung konsep restoran yang dipadukan dengan berbagai *spot* foto *instagramable* serta menawarkan suasana alam yang sejuk dengan pepohonan yang rindang. HeHa *Forest* resmi dibuka pada 17 April 2023 dan berlokasi di Jalan Kaliurang KM 22, Banteng, Hargobinangun, Pakem, Yogyakarta. Sejak dibuka, destinasi ini berhasil menarik perhatian wisatawan, khususnya wisatawan lokal. Selain menyajikan pemandangan alam, HeHa *Forest* juga menyediakan berbagai fasilitas seperti *Bunny Zone*, restoran, *food and beverage stall*, ruang pertemuan, dan *souvenir shop*. Daya tarik utama destinasi ini terletak pada konsep wisata kekinian yang memadukan nuansa alam dengan berbagai *spot* foto yang sesuai dengan tren media sosial (Dwi Ayunda, 2024). Beragam fasilitas dan konsep wisata tersebut membentuk ekspektasi tertentu bagi pengunjung sebelum melakukan kunjungan. Oleh karena itu, diperlukan analisis terhadap opini pengunjung untuk mengetahui apakah pengalaman yang diperoleh telah sesuai dengan ekspektasi tersebut.

Analisis sentimen atau *opinion mining* pada sektor pariwisata merupakan pendekatan yang digunakan untuk memahami opini, penilaian, sikap, dan emosi pengunjung terhadap suatu destinasi wisata melalui data berbasis teks. Analisis sentimen memungkinkan identifikasi kecenderungan opini menjadi positif, netral, maupun negatif sehingga dapat digunakan untuk memproyeksikan persepsi masyarakat terhadap suatu objek wisata. Namun, analisis sentimen hanya menunjukkan kecenderungan polaritas opini dan belum mampu menjelaskan topik atau aspek yang sedang dibahas oleh pengunjung. Oleh karena itu, diperlukan metode *topic modelling* untuk mengelompokkan komentar berdasarkan tema pembahasannya.

Topic modelling merupakan teknik yang digunakan untuk mengidentifikasi tema tersembunyi dari kumpulan data teks dalam jumlah besar. Teknik ini mampu mengenali pola dan mengelompokkan dokumen berdasarkan topik yang memiliki kemunculan kata yang serupa (Jeong, 2025). Salah satu pendekatan *topic modelling* yang banyak digunakan adalah *Latent Dirichlet Allocation* (LDA). LDA merupakan metode yang terbukti efektif dalam mengidentifikasi topik karena setiap topik direpresentasikan

sebagai kumpulan kata yang sering muncul secara bersamaan dalam keseluruhan dokumen (Mayasari dkk., 2021). Selain mampu mengidentifikasi topik dominan, LDA juga banyak dimanfaatkan untuk menganalisis tren penelitian maupun tema yang berkembang (Husnul dkk., 2025). Meskipun demikian, hasil LDA hanya menghasilkan kumpulan topik tanpa menunjukkan tingkat prioritas permasalahan yang perlu ditangani terlebih dahulu. Oleh karena itu, diperlukan metode pengambilan keputusan, yaitu TOPSIS, untuk menentukan prioritas aspek yang memerlukan perbaikan.

Metode *Technique for Order Preference by Similarity to Ideal Solution* (TOPSIS) digunakan untuk menentukan prioritas aspek yang perlu diperbaiki berdasarkan hasil analisis topik dan sentimen. Metode ini memilih alternatif terbaik berdasarkan kedekatan terhadap solusi ideal positif dan jarak terjauh dari solusi ideal negatif. Solusi ideal positif merupakan alternatif yang memberikan nilai keuntungan (*benefit*) tertinggi dan biaya (*cost*) terendah, sedangkan solusi ideal negatif merupakan alternatif dengan keuntungan terendah dan biaya tertinggi. Dengan demikian, TOPSIS mampu memberikan urutan prioritas permasalahan yang perlu menjadi fokus perbaikan sehingga proses pengambilan keputusan menjadi lebih objektif (Kristina, 2018).

Setelah prioritas permasalahan diperoleh menggunakan TOPSIS, diperlukan suatu pendekatan yang mampu mengembangkan alternatif solusi secara sistematis. Salah satu teknik yang dapat digunakan adalah *Lotus Blossom*, yaitu teknik pengembangan ide yang diawali dari satu permasalahan utama kemudian dikembangkan menjadi delapan gagasan utama dan selanjutnya diperluas kembali menjadi ide-ide yang lebih spesifik. Pendekatan ini membantu menghasilkan rekomendasi yang terstruktur melalui proses eksplorasi berbagai kemungkinan solusi sehingga lebih mudah diterapkan dalam penyusunan strategi perbaikan. *Lotus Blossom* telah banyak digunakan sebagai teknik berpikir kreatif dan pemecahan masalah karena mampu mengorganisasi ide secara sistematis dan mendorong munculnya solusi yang lebih komprehensif (Kleinpaul dkk., 2025).

Oleh karena itu, penelitian ini diharapkan dapat mengoptimalkan HeHa *Forest* berdasarkan opini dan persepsi yang ada dari pengunjung melalui pemanfaatan LDA dan TOPSIS. Hasil dari penelitian inipun juga dapat dimanfaatkan bagi pemangku kepentingan pariwisata agar dapat merancang strategi yang optimal dalam pengembangan wisata yang berkelanjutan sehingga berdampak pada kepuasan pengunjung.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, maka didapatkan rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Apa saja topik utama yang muncul dalam ulasan pengunjung terhadap HeHa *Forest*?
2. Bagaimana kecenderungan sentimen pengunjung terhadap HeHa *Forest* berdasarkan ulasan yang diberikan?
3. Bagaimana menentukan prioritas aspek layanan yang perlu ditingkatkan berdasarkan hasil analisis topik dan sentimen?
4. Bagaimana rekomendasi strategi pengembangan wisata bagi pengelola HeHa *Forest* berdasarkan prioritas permasalahan yang diperoleh?

1.3 Tujuan Penelitian

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

1. Mengidentifikasi topik utama yang muncul dalam ulasan pengunjung terhadap HeHa *Forest*
2. Menganalisis kecenderungan sentimen pengunjung terhadap HeHa *Forest* berdasarkan ulasan yang diberikan
3. Menentukan prioritas aspek layanan yang perlu ditingkatkan berdasarkan hasil analisis topik dan *sentiment*
4. Merumuskan rekomendasi strategi pengembangan wisata bagi pengelola HeHa *Forest* berdasarkan prioritas permasalahan yang diperoleh.

1.4 Manfaat Penelitian

Berikut merupakan manfaat dari adanya penelitian ini:

1. Bagi peneliti
 - a. Meningkatkan keterampilan dan wawasan peneliti terhadap penerapan konsep analisis sentimen *Latent Dirichlet Allocation* (LDA), dan *Technique For Orders Reference by Similarity to Ideal Solution* (TOPSIS).
 - b. Menambah wawasan tentang penerapan keilmuan dalam bidang Teknik Industri.
2. Bagi Universitas
 - a. Menjadi referensi tambahan bagi civitas akademika dalam pengembangan penelitian yang berkaitan dengan analisis data dan pengambilan keputusan.

- b. Mendukung pengembangan keilmuan di bidang Teknik Industri, khususnya dalam penerapan analisis ulasan dan evaluasi layanan.
- c. Mendorong pemanfaatan data digital sebagai sumber informasi dalam penelitian akademik.

3. Bagi HeHa *Forest*

- a. Memberikan informasi mengenai persepsi dan pengalaman pengunjung berdasarkan ulasan yang tersedia secara daring.
- b. Menyediakan gambaran mengenai aspek layanan yang menjadi perhatian utama pengunjung.
- c. Memberikan rekomendasi prioritas perbaikan dan strategi pengembangan yang dapat digunakan sebagai dasar dalam pengambilan keputusan untuk meningkatkan kualitas layanan dan daya tarik wisata.

1.5 Batasan Penelitian

Berikut ini merupakan batasan penelitian:

1. Penelitian ini menggunakan data ulasan pengunjung yang diperoleh dari *platform Google Maps* dan *Instagram* pada periode tahun 2023 hingga Februari 2026, dengan objek penelitian difokuskan pada HeHa *Forest* yang berlokasi di Jalan Kaliurang, Sleman, Yogyakarta.
2. Penelitian ini berfokus pada analisis ulasan pengunjung untuk mengidentifikasi topik, menganalisis sentimen, serta menentukan prioritas perbaikan layanan dan rekomendasi strategi pengembangan, tanpa membahas aspek finansial maupun faktor eksternal lainnya secara mendalam.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Literatur

2.1.1 Web Scrapping dalam Data Pariwisata

Penggunaan *web scraping* sebagai metode pengumpulan data wisata berbasis ulasan, media sosial, dan jejak digital pengguna telah meningkat sebagai akibat dari perkembangan *platform* digital. Blanco-Moreno dkk. (2024a) menemukan bahwa *scraping* 27.088 unggahan *Instagram* dapat digunakan untuk menganalisis keterlibatan destinasi wisata. Di sisi lain, Jiang dkk. (2025) menambang puluhan ribu ulasan wisata untuk menilai elemen daya tarik buku perjalanan. Hasil ini menunjukkan bahwa *scraping web* telah berkembang dari sekadar alat pengambilan data menjadi dasar penelitian pariwisata berbasis data besar.

Metode serupa juga digunakan untuk mengevaluasi persepsi wisatawan dari *platform* ulasan dalam hal destinasi alam dan ulasan pengunjung. Pada sisi lain, Ramdani & Cahyana (2025) mengumpulkan 2.469 ulasan wisata budaya melalui *scraping Google Maps* dan *Twitter*. Ini menunjukkan bahwa *scraping* berlaku untuk objek wisata yang sangat bergantung pada persepsi publik, seperti *HeHa Forest*.

2.1.2 Analisis Sentimen dalam Penelitian Pariwisata

Analisis sentimen telah banyak digunakan untuk mengetahui bagaimana pengunjung melihat layanan, pengalaman, dan fitur destinasi. Mensah dkk. (2025) menunjukkan bahwa kepuasan wisatawan lintas destinasi Afrika dapat diukur melalui ulasan pengguna. Pada sisi Bigne dkk. (2023) menunjukkan hubungan yang kuat antara isi ulasan pengguna dan rating bintang pada destinasi wisata. Artinya, sentimen bukan hanya melabelkan hal-hal yang baik atau buruk, tetapi juga membantu menjelaskan bagaimana pengalaman wisata diterjemahkan ke dalam nilai numerik.

Metode ini juga semakin berkembang di tingkat nasional. Singgalen (2023) menggunakan analisis sentimen pada ulasan Pulau Komodo dan Pulau Rinca di *TripAdvisor* Wulan Sari dkk. (2023) membandingkan *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) untuk ulasan wisata dan Sa'adah dkk. (2025) menggunakan beberapa algoritma pembelajaran yang diawasi untuk mengidentifikasi perbedaan pendapat wisatawan Indonesia dari ulasan *Google Maps*. Secara keseluruhan,

studi-studi ini menegaskan bahwa analisis sentimen adalah teknik yang efektif untuk mendapatkan suara wisatawan dari teks ulasan berbahasa Indonesia.

2.1.3 Topic Modelling Menggunakan Latent Dirichlet Allocation (LDA)

Riset pariwisata sering menggunakan LDA untuk menemukan tema utama dalam kumpulan ulasan yang besar. Kirilenko & Stepchenkova (2025) membandingkan teknologi *topic modelling* baru dengan LDA pada data khas pariwisata seperti *review titles*, *TripAdvisor*, dan komentar *online*. Hasilnya menunjukkan bahwa LDA masih merupakan dasar penting untuk ekstraksi topik pariwisata. Selain itu, Wu dkk. (2024) melakukan peramalan permintaan hotel yang lebih akurat dengan menggabungkan LDA dan sentimen. Ini menunjukkan bahwa informasi topik lebih bermanfaat daripada hanya melihat polaritas opini.

Merawati dkk. (2021) telah menggabungkan analisis sentimen dan LDA untuk pariwisata Lombok pada konteks nasional. Studi ini penting karena menunjukkan bahwa ulasan wisata Indonesia dapat dilihat dari perspektif emosi dan tema diskusi, sehingga pengelola dapat mengidentifikasi masalah yang paling umum.

2.1.4 Metode Pengambilan Keputusan TOPSIS

Dalam riset pariwisata, metode multikriteria seperti TOPSIS banyak digunakan untuk mengubah temuan teks menjadi prioritas tindakan. Wang dkk. (2024) mengevaluasi kualitas layanan budaya dan pariwisata Huizhou dengan entropy-weight TOPSIS, sedangkan Akten (2025) menggunakan SWOT-TOPSIS untuk menilai strategi pariwisata berkelanjutan. Kedua penelitian ini menunjukkan bahwa TOPSIS cocok ketika peneliti perlu mengurutkan alternatif kuantitatif berdasarkan berbagai kriteria sekaligus.

Selain itu, penggunaan TOPSIS muncul dalam situasi yang lebih khusus. Suárez-Rojas dkk. (2025) menggunakan *fuzzy-hybrid* TOPSIS untuk menjelaskan perilaku penonton ikan paus, Chen dkk. (2024) menggunakan *FAHP-fuzzy* TOPSIS untuk rekomendasi hotel, dan Wang dkk. (2024) menunjukkan bahwa metode TOPSIS dapat digunakan untuk mengevaluasi kualitas layanan pariwisata secara kuantitatif berdasarkan berbagai kriteria. Ini menunjukkan bahwa TOPSIS sangat fleksibel untuk digunakan pada tingkat prioritas layanan dan strategi.

2.1.5 Integrasi Metode dalam Penelitian Wisata

Menurut arah penelitian terbaru, ada kecenderungan untuk mengintegrasikan metode Wu dkk. (2024) menggunakan model topik dan analisis perasaan untuk memprediksi permintaan hotel, sementara Jiang dkk. (2025) menggunakan *review mining* yang

menggabungkan *extraction* topik dan analisis perasaan untuk mengukur elemen atraktif. Integrasi seperti ini penting karena *topic modelling* hanya menunjukkan topik diskusi, sedangkan sentimen hanya menunjukkan kecenderungan pendapat.

Sebaliknya, penelitian TOPSIS menunjukkan bahwa hasil analisis yang sudah terstruktur masih dapat digunakan sebagai dasar keputusan prioritas. Oleh karena itu, ada ruang metodologis yang kuat untuk menghubungkan hasil topik, skor sentimen, dan kemudian peringkat prioritas dalam satu alur analisis.

2.1.6 Research Gap dan Posisi Penelitian

Penelitian tentang pariwisata berbasis data ulasan daring telah berkembang berkat penggunaan teknik scraping web untuk mengumpulkan data dari berbagai *platform* digital serta penggunaan analisis sentimen dan *topic modelling* untuk memahami persepsi wisatawan (Blanco-Moreno dkk., 2024a; Jiang dkk., 2025; Kirilenko & Stepchenkova, 2025; Mensah dkk., 2025). Dalam evaluasi layanan pariwisata, teknik pengambilan keputusan seperti TOPSIS juga telah digunakan untuk menentukan prioritas alternatif berdasarkan berbagai kriteria (Akten, 2025; Wang dkk., 2024). Meskipun demikian, sebagian besar penelitian masih menggunakan metode ini secara terpisah, sehingga hasil yang diperoleh cenderung bersifat parsial dalam memahami pendapat wisatawan dan menentukan perbaikan layanan mana yang paling penting.

Selain itu, penelitian yang menggabungkan analisis sentimen dan *topic modelling* belum banyak berfokus pada menentukan prioritas perbaikan layanan secara operasional dan hanya melakukan analisis eksploratif (Wu dkk., 2024). Selain itu, penggunaan data ulasan *online* sebagai dasar pengambilan keputusan belum ideal, terutama ketika metode pengambilan keputusan multikriteria digabungkan. Oleh karena itu, penelitian ini menyarankan metode terintegrasi yang menggabungkan analisis sentimen berbasis *InSet Lexicon*, pembagian topik menggunakan LDA, dan metode TOPSIS ke dalam satu kerangka analisis berbasis data ulasan yang menggunakan *Instagram* dan *Google Maps*. Diharapkan bahwa metode ini akan memungkinkan pengelola wisata HeHa Forest untuk mendapatkan analisis yang lebih menyeluruh dan menetapkan langkah-langkah kuantitatif untuk meningkatkan layanan mereka.

Tabel 2. 1 Posisi Penelitian

No	Penulis	Sumber Data (Jenis Data)		Metode			Objek
		Tradisional (Wawancara/Survei)	Online Review Platform	Sentiment Analysis	LDA	TOPSIS	
1	Merawati dkk. (2021)		✓	✓	✓		Lombok
2	Singgale (2023)		✓	✓			Komodo
3	Wulan dkk. (2023)		✓	✓			Wisata lokal
4	Bigne dkk. (2023)		✓	✓			Destinasi wisata
5	Chen dkk. (2024)	✓				✓	<i>Sustainable tourism</i>
6	Wu dkk. (2024)		✓	✓	✓		Hotel
7	Wang dkk. (2024)	✓				✓	Pariwisata budaya
8	Blanco-Moreno dkk. (2024)		✓	✓	✓		Destinasi wisata
9	Kirilenko (2025)	✓				✓	Hotel
10	Suárez-Rojas dkk. (2025)	✓				✓	Wisata bahari
11	Akten (2025)		✓		✓		Pariwisata umum
12	Jiang dkk. (2025)		✓	✓	✓		<i>Library tourism</i>
13	Mensah dkk. (2025)		✓	✓			Destinasi Afrika
14	Ramdani (2025)		✓	✓	✓	✓	Wisata budaya
15	Miftahus Sa'adah dkk. (2025)		✓	✓			Destinasi Indonesia
16	Penelitian Ini		✓	✓	✓	✓	HeHa Forest

2.2 Landasan Teori

2.2.1 Pariwisata

Salah satu sektor yang memainkan peran penting dalam pertumbuhan ekonomi dan pembangunan sebuah wilayah adalah pariwisata, khususnya dalam hal meningkatkan pendapatan dan menciptakan peluang kerja. Selain itu, pariwisata juga terkait dengan aktivitas perjalanan, yang mencakup pengalaman dan interaksi pengunjung dengan lokasi tersebut. Oleh karena itu, pengalaman pengunjung yang berkualitas tinggi merupakan salah satu komponen penting dalam keberhasilan destinasi wisata (Mensah dkk., 2025)

Sekarang pariwisata tidak hanya dilihat sebagai aktivitas fisik, tetapi juga sebagai pengalaman yang dipengaruhi oleh persepsi dan penilaian wisatawan; persepsi ini dapat mencerminkan tingkat kepuasan wisatawan dan kualitas layanan yang diberikan oleh pengelola destinasi. Oleh karena itu, memahami persepsi wisatawan sangat penting untuk meningkatkan kualitas layanan dan daya saing destinasi wisata.

2.2.2 Web Scraping

Web Scraping adalah metode pengambilan data otomatis dari halaman *web* untuk mendapatkan informasi yang sangat besar yang dapat digunakan untuk proses analisis lanjutan. Teknik ini memungkinkan pengumpulan data yang efektif tanpa harus dilakukan secara manual, sehingga sangat relevan untuk penelitian yang memanfaatkan data digital yang sangat besar (Blanco-Moreno dkk., 2024a).

Dalam industri pariwisata, *web scraping* digunakan untuk mengumpulkan data ulasan pelanggan dari berbagai *platform digital* seperti media sosial, *Google Maps*, dan *TripAdvisor*. Pengalaman dan persepsi wisatawan terhadap suatu destinasi biasanya diwakili oleh teks ulasan, komentar, dan rating. Menurut Jiang dkk. (2025), *teknik text mining* dapat digunakan untuk mengidentifikasi elemen daya tarik destinasi dari data yang dikumpulkan dari *platform* ulasan.

Tetapi data yang dikumpulkan melalui *web scraping* biasanya tidak terstruktur dan memerlukan proses pengolahan tambahan untuk dianalisis secara sistematis. Oleh karena itu, metode *web scraping* digunakan sebagai tahap awal untuk mengumpulkan data ulasan wisatawan dari *Instagram* dan *Google Maps*. Data ini kemudian akan diolah menggunakan teknik *text mining*, analisis sentimen, dan *topic modelling* untuk menghasilkan informasi.

2.2.3 Text Mining

Text mining adalah teknik yang digunakan untuk mengekstraksi informasi dari data teks yang tidak terstruktur menjadi informasi yang dapat dianalisis dan digunakan. Dalam penelitian berbasis data ulasan daring, ini digunakan untuk mengolah teks ulasan wisatawan untuk menemukan pola, kecenderungan, dan informasi penting yang terkandung di dalamnya (Kirilenko, 2025). Karena data ulasan yang diperoleh dari *platform* digital biasanya berbentuk teks bebas yang tidak dapat langsung dianalisis secara kuantitatif, proses ini menjadi penting.

Untuk mempersiapkan data sebelum dianalisis lebih lanjut, *text mining* biasanya melibatkan proses *pre-processing* seperti *stopword removal* dimana penghapusan kata yang dianggap *noise*, selanjutnya pada tahap *case folding* dengan mengecilkan setiap kata, dilanjut tahap tokenisasi dengan memisahkan setiap kata dengan tanda baca seperti [*sample*], setelah hal tersebut dilanjut dengan *filtering* dengan menghilangkan tanda baca, *mention*, dan emotikon, setelah itu dilanjut dengan proses *stemmed* tujuan dari tahap ini adalah untuk menyatukan berbagai turunan kata yang memiliki dasar yang sama,

sehingga dapat dianggap sebagai satu entitas dalam analisis, setelah melewati tahapan-tahapan pada *pre-processing*, token-token tersebut disatukan Kembali menjadi sebuah kalimat yang bertujuan untuk mempermudah proses lanjutan. Dalam industri pariwisata, penggunaan *text mining* pada data ulasan *online* dapat meningkatkan pemahaman tentang pengalaman dan preferensi pengunjung (Wu dkk., 2024). Oleh karena itu, *text mining* digunakan sebagai tahap awal pengolahan data sebelum *topic modelling* dan analisis sentimen dilakukan untuk menghasilkan informasi yang lebih lengkap.

2.2.4 User Generated Content

Konten yang dihasilkan oleh pengguna pada *platform* digital, seperti ulasan, komentar, *rating*, dan unggahan di media sosial, disebut *User-Generated Content* (UGC). Dalam industri pariwisata, UGC sangat penting karena mencerminkan pengalaman dan persepsi wisatawan secara langsung terhadap suatu destinasi (Mensah dkk., 2025). Konten yang dibuat oleh pengguna biasanya spontan dan tidak terstruktur, sehingga lebih realistis daripada data yang diperoleh melalui metode konvensional.

UGC banyak digunakan dalam penelitian pariwisata seiring dengan meningkatnya penggunaan *platform* digital untuk menganalisis kepuasan, preferensi, dan pengalaman wisatawan. Blanco-Moreno dkk. (2024a) menemukan bahwa interaksi pengguna di media sosial dapat digunakan untuk mengetahui keterlibatan dan persepsi pengguna terhadap destinasi wisata. Oleh karena itu, sebagai sumber data utama untuk penelitian ini, UGC digunakan dari *Google Maps* dan *Instagram*. Data ini kemudian diolah menggunakan teknik *text mining*, analisis sentimen, dan *topic modelling* untuk mendapatkan informasi yang lebih lengkap.

2.2.6 InSet Lexicon

InSet Lexicon, juga dikenal sebagai Indonesia *Sentiment Lexicon*, adalah kamus sentimen yang dibuat khusus untuk bahasa Indonesia. Itu memberikan nilai polaritas terhadap kata-kata yang memiliki makna baik positif maupun negatif. Metode analisis sentimen berbasis leksikon memanfaatkan skor kamus kata untuk menentukan sentimen teks. Leksikon lokal dapat menangkap fitur bahasa, konteks, dan variasi ekspresi yang tidak selalu dapat diakomodasi oleh leksikon berbahasa Inggris, yang membuat penggunaan leksikon lokal sangat penting (Kurniawan, 2018).

Untuk meningkatkan analisis sentimen pada teks berbahasa Indonesia, khususnya data informal seperti ulasan dan komentar di media sosial, dibuat leksikon sentimen bahasa Indonesia. Karena konteks lokal dan struktur bahasa, penggunaan leksikon khusus bahasa

Indonesia dapat memberikan hasil klasifikasi sentimen yang lebih representatif, menurut penelitian (Kurniawan, 2018).

Metode berbasis leksikon memiliki beberapa keunggulan dibandingkan dengan pendekatan berbasis *machine learning*. Ini termasuk tidak memerlukan data latih berlabel, lebih mudah digunakan, dan lebih efisien dalam pengolahan data. Selain itu, metode ini lebih cocok untuk data ulasan wisata yang menggunakan bahasa sehari-hari dan tidak terstruktur (Wulan dkk., 2023). Namun, pendekatan berbasis leksikon tidak dapat memahami kalimat yang kompleks karena sarkasme atau ambiguitas makna.

Untuk melakukan analisis sentimen, ada sejumlah pendekatan yang dapat digunakan, termasuk metode berbasis *machine learning*, metode berbasis leksikon, dan metode *hybrid*. Setiap pendekatan memiliki karakteristik yang berbeda dalam menentukan polaritas sentimen teks, sehingga diperlukan perbandingan antar pendekatan untuk menentukan yang paling sesuai dengan data dan kebutuhan penelitian (Rahmaningtyas, 2017). Berikut merupakan tabel perbandingan metode analisis sentimen:

Tabel 2. 2 Tabel perbandingan metode analisis sentimen

Metode	Bahasa Indonesia	Tanpa Data Training	Cocok untuk Ulasan Wisata	Dapat Mengolah Bahasa Informal	Kompleksitas Rendah
<i>InSet</i>	✓	✓	✓	✓	✓
<i>Lexicon</i>	✓	✓	✓	✓	✓
VADER	✗	✓	✓	✓	✓
<i>Hybrid</i>	✓	✗	✓	✓	✗
<i>Machine Learning</i>	✓	✗	✓	✓	✗

Perbandingan tersebut menunjukkan bahwa *InSet Lexicon* lebih cocok untuk analisis sentimen pada data teks berbahasa Indonesia, terutama ulasan wisatawan yang tidak terstruktur. Dibandingkan dengan metode berbasis *machine learning* yang memerlukan data *training*, serta metode seperti VADER yang tidak dirancang untuk Bahasa Indonesia, *InSet Lexicon* lebih efektif dan relevan dalam mengidentifikasi polaritas kata dalam konteks lokal. Dalam penelitian ini, metode ini dianggap cukup untuk memberikan gambaran umum tentang perasaan wisatawan, meskipun memiliki keterbatasan dalam memahami konteks yang kompleks (Rahmaningtyas, 2017).

Dalam penelitian ini, *InSet Lexicon* digunakan untuk mengidentifikasi polaritas sentimen dari ulasan wisatawan terhadap destinasi HeHa Forest. Hasil analisis sentimen ini kemudian digunakan sebagai dasar untuk analisis lebih lanjut, yang melibatkan

penggabungan dengan *topic modelling* menggunakan LDA dan metode pengambilan keputusan TOPSIS. Ini membuat hasil analisis menjadi lebih komprehensif.

2.2.5 Analisis Sentimen

Analisis sentimen adalah teknik *text mining* yang digunakan untuk mengidentifikasi opini, emosi, dan sikap yang terkandung dalam suatu teks. Dalam industri pariwisata, analisis sentimen digunakan untuk memahami persepsi wisatawan terhadap suatu destinasi berdasarkan ulasan yang diberikan di platform digital. Melalui analisis ini, teks ulasan dapat dikategorikan ke dalam kategori tertentu, seperti positif, negatif, atau netral, sehingga membantu memahami persepsi wisatawan terhadap destinasi (Mensah dkk., 2025).

Dua pendekatan umum untuk analisis sentimen adalah berbasis *machine learning* dan berbasis leksikon. Pendekatan pertama menggunakan data latih berlabel untuk membangun model klasifikasi, sedangkan pendekatan kedua menggunakan kamus kata yang memiliki nilai polaritas untuk mengidentifikasi sentimen teks. Kedua metode ini telah banyak digunakan dalam penelitian pariwisata untuk menganalisis ulasan wisatawan dan menemukan kecenderungan pendapat mereka tentang suatu destinasi (Wulan dkk., 2023).

Dalam penelitian ini, analisis sentimen digunakan sebagai dasar untuk mengukur kecenderungan opini wisatawan terhadap destinasi HeHa *Forest* berdasarkan data ulasan yang diperoleh dari *platform* digital. Hasil dari analisis sentimen ini selanjutnya digunakan sebagai bagian dari proses analisis lanjutan, khususnya ketika digabungkan dengan metode *topic modelling* dan pengambilan keputusan untuk memberikan gambaran yang lebih baik.

2.2.7 Topic Modelling

Topic modelling adalah teknik *text mining* yang digunakan untuk menemukan pola atau topik tersembunyi dalam kumpulan ulasan. Teknik ini memungkinkan peneliti untuk menemukan struktur tematik dari data teks yang tidak terstruktur tanpa memerlukan label awal, sehingga sangat relevan untuk analisis data ulasan wisatawan yang sangat beragam (Kirilenko, 2025).

Topic modelling digunakan dalam industri pariwisata untuk menentukan topik utama yang sering dibahas oleh wisatawan dalam ulasan mereka, seperti fasilitas, pelayanan, harga, dan pengalaman secara keseluruhan. Wu dkk. (2024) mengatakan bahwa data

ulasan *online* dapat membantu kita memahami preferensi wisatawan dan faktor-faktor yang memengaruhi kepuasan mereka dengan destinasi.

Topic modelling memiliki keunggulan dalam menghasilkan representasi topik yang sistematis dan memungkinkan pengolahan data teks yang sangat besar. Oleh karena itu, pendekatan ini dapat membantu pengelola destinasi wisata dalam menentukan aspek penting dari pengalaman wisatawan yang harus diperhatikan. Oleh karena itu, topik *modelling* digunakan dalam penelitian ini untuk mengidentifikasi topik utama dari ulasan wisatawan. Topik-topik ini akan dianalisis lebih lanjut dengan menggunakan metode *Latent Dirichlet Allocation* (LDA).

2.2.8 *Latent Dirichlet Allocation* (LDA)

Latent Dirichlet Allocation (LDA) digunakan untuk menemukan topik tersembunyi dalam kumpulan data teks atau ulasan, metode ini bekerja dengan melihat pola kemunculan kata yang sering muncul secara bersamaan dalam ulasan. Dalam LDA, setiap ulasan dapat memiliki lebih dari satu topik, tetapi setiap topik terdiri dari beberapa kata yang saling terkait. Sehingga, LDA dapat membantu memecah ulasan wisatawan yang tidak terstruktur menjadi beberapa topik utama yang lebih mudah dipahami dan dianalisis. (Kirilenko, 2025).

Pada penelitian ini, LDA digunakan untuk mengidentifikasi topik utama yang muncul dalam ulasan pengunjung HeHa *Forest*. Dengan cara ini, komentar dan ulasan pengunjung dapat dikelompokkan berdasarkan pola kata dominan, seperti fasilitas, harga, layanan, makanan, *Spot* foto, dan pengalaman wisata. Hasil pemodelan topik ini kemudian digunakan sebagai dasar untuk proses interpretasi topik, analisis sentimen, dan penentuan prioritas perbaikan layanan menggunakan TOPSIS.

LDA telah banyak digunakan dalam penelitian pariwisata untuk mengidentifikasi tema utama dalam ulasan wisatawan, seperti kualitas layanan, fasilitas, dan pengalaman wisata. Penelitian oleh Wu dkk. (2024) menggunakan LDA untuk mengidentifikasi topik utama pada ulasan wisatawan terhadap destinasi wisata dan menghasilkan informasi terkait opini pengunjung. Selain itu, penelitian oleh Kirilenko, (2025) ini menunjukkan bahwa LDA dapat digunakan untuk mengelompokkan topik berdasarkan ulasan wisatawan. Berikut merupakan tabel perbandingan *topic modelling*:

Tabel 2. 3 Tabel perbandingan *topic modelling*

Metode	Cocok untuk Data Tidak Terstruktur	Mudah Diinterpretasikan	Menghasilkan Topik Representatif	Cocok untuk Bahasa Indonesia	<i>Coherence Score</i> Lebih Baik
--------	------------------------------------	-------------------------	----------------------------------	------------------------------	-----------------------------------

LDA	✓	✓	✓	✓	✓
BERT opic	✓	✓	✓	✓	X

Dalam penelitian ini, *Latent Dirichlet Allocation* (LDA) dipilih karena mampu mengelompokkan ulasan berdasarkan pola kemunculan kata yang sering muncul secara bersamaan. Dengan metode ini, data ulasan yang tidak terstruktur dapat dibagi menjadi beberapa topik utama sehingga lebih mudah dipahami dan dianalisis. LDA memiliki kemampuan yang lebih baik untuk mengidentifikasi keterkaitan antar kata dalam suatu topik secara alami dibandingkan dengan metode seperti *BERTopic*, yang berbasis *clustering numerik*. Namun, metode ini dapat dioptimalkan dengan skor *coherence* sehingga menghasilkan jumlah topik yang sesuai dengan karakteristik data. Akibatnya, metode ini dianggap lebih relevan untuk digunakan dalam penelitian ini (Tiwari dkk., 2023).

Metode LDA digunakan untuk menentukan topik utama yang muncul dalam ulasan wisatawan tentang destinasi *HeHa Forest*. Hasil pemodelan topik ini kemudian digunakan sebagai dasar untuk pengolahan lebih lanjut, di mana hasil analisis sentimen berbasis *InSet Lexicon* dan metode pengambilan keputusan TOPSIS digunakan. Dengan demikian, penelitian ini dapat menghasilkan analisis yang lebih menyeluruh dan berguna.

2.2.9 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

Metode pengambilan keputusan multikriteria yang dikenal sebagai *Technique for Order Preference by Similarity to Ideal Solution* (TOPSIS) digunakan untuk memilih alternatif terbaik berdasarkan seberapa dekatnya dengan solusi ideal positif dan seberapa jauhnya dari solusi ideal negatif. Metode TOPSIS pertama kali diperkenalkan oleh Hwang dan Yoon pada tahun 1981 sebagai salah satu metode pengambilan keputusan multikriteria yang mempertimbangkan kedekatan terhadap solusi ideal positif dan negatif. Metode ini juga banyak digunakan dalam berbagai industri, seperti pariwisata (Akten, 2025).

Konsep dasar TOPSIS adalah bahwa pilihan terbaik merupakan alternatif yang memiliki jarak paling dekat dengan solusi ideal positif dan paling jauh dari solusi ideal negatif. Solusi ideal positif merupakan kondisi terbaik untuk setiap kriteria, sedangkan solusi ideal negatif merupakan kondisi terburuk dari setiap kriteria. Dengan konsep tersebut, TOPSIS mampu menghasilkan proses pengambilan keputusan yang lebih objektif karena mempertimbangkan kedekatan setiap alternatif terhadap kedua solusi ideal secara bersamaan (Akten, 2025). Hasil akhir metode TOPSIS berupa peringkat

alternatif berdasarkan nilai kedekatan relatif (*closeness coefficient*), sehingga alternatif dengan nilai preferensi tertinggi dipilih sebagai prioritas utama dalam pengambilan keputusan (Mao dkk., 2021). Oleh karena itu, pada penelitian ini hanya menggunakan satu topik prioritas pada setiap *platform*, yaitu topik dengan nilai preferensi TOPSIS tertinggi, karena topik tersebut merepresentasikan alternatif yang paling layak diprioritaskan untuk diperbaiki. Dalam penerapannya, metode TOPSIS dilakukan melalui beberapa tahapan, yaitu sebagai berikut.

1. Normalisasi Matriks Keputusan

Berikut merupakan rumus Normalisasi Matriks:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}^2} \quad (2.1)$$

Keterangan:

r_{ij} = nilai normalisasi

x_{ij} = nilai alternatif ke-i pada kriteria ke-j

m = jumlah alternatif

2. Pembentukan Matriks Normalisasi Terbobot

Berikut merupakan rumus pembentukan matriks normalisasi terbobot:

$$v_{ij} = w_{ij} \times r_{ij} \quad (2.2)$$

Keterangan:

v_{ij} = nilai terbobot

w_{ij} = bobot kriteria

3. Menentukan Solusi Ideal Positif dan Negatif

Berikut merupakan rumus solusi ideal positif dan *negative*:

$$A^+ = (v_1^+, v_2^+, \dots, v_n^+) \quad (2.3)$$

$$A^- = (v_1^-, v_2^-, \dots, v_n^-) \quad (2.4)$$

Keterangan:

$v_1^+ = \max (v_{ij})$ nilai solusi ideal positif pada kriteria ke-j

$v_1^- = \min (v_{ij})$ nilai solusi ideal negatif pada kriteria ke-j

4. Menghitung Jarak Terhadap Solusi Ideal

Jarak antara alternatif A_i dengan solusi ideal positif dirumuskan sebagai:

$$D_i^+ = \sqrt{\sum_{j=1}^n (y_i^+ - y_{ij})^2} \quad (2.5)$$

Jarak antara alternatif A_i dengan solusi ideal negatif dirumuskan sebagai:

$$D_i^- = \sqrt{\sum_{j=1}^n (y_i^- - y_{ij})^2} \quad (2.6)$$

Keterangan:

D_i^+ = jarak alternatif ke- i terhadap solusi ideal positif

D_i^- = jarak alternatif ke- i terhadap solusi ideal negatif

5. Menghitung Nilai Prefensi

Berikut merupakan rumus menghitung nilai prefensi:

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad (2.7)$$

Keterangan:

C_i = nilai prefensi alternatif

Nilai C_i berada pada rentang 0 hingga 1, dimana semakin besar nilai C_i , maka alternatif tersebut semakin mendekati solusi ideal dan menjadi prioritas yang lebih baik (Akten, 2025).

2.2.10 Integrasi Metode

Untuk mendapatkan hasil yang lebih komprehensif, penelitian ini menggunakan pendekatan terintegrasi yang menggabungkan berbagai teknik analisis. Analisis dimulai dengan data UGC yang dikumpulkan melalui *web scraping* dari *platform* seperti *Instagram* dan *Google Maps*. Data ini kemudian diproses melalui teknik *text mining* (Blanco-Moreno dkk., 2024a).

Selain itu, kecenderungan opini wisatawan ditentukan melalui analisis sentimen berbasis *InSet Lexicon*. Di sisi lain, topik utama dalam ulasan ditentukan melalui metode LDA (Kirilenko, 2025). Hasil dari kedua pendekatan ini kemudian digunakan sebagai dasar untuk menggunakan metode TOPSIS untuk menentukan elemen apa yang harus diperbaiki dengan prioritas tertinggi (Akten, 2025). Pendekatan ini memungkinkan penelitian untuk menemukan topik dan pendapat serta membuat rekomendasi kuantitatif untuk perbaikan layanan yang paling penting.

BAB III

METODE PENELITIAN

3.1 Subjek Penelitian

Penelitian ini berfokus pada persepsi dan perasaan pengunjung terhadap destinasi wisata HeHa *Forest*. Ulasan daring digunakan untuk menentukan topik utama dan menentukan aspek layanan mana yang harus diperbaiki.

3.2 Objek Penelitian

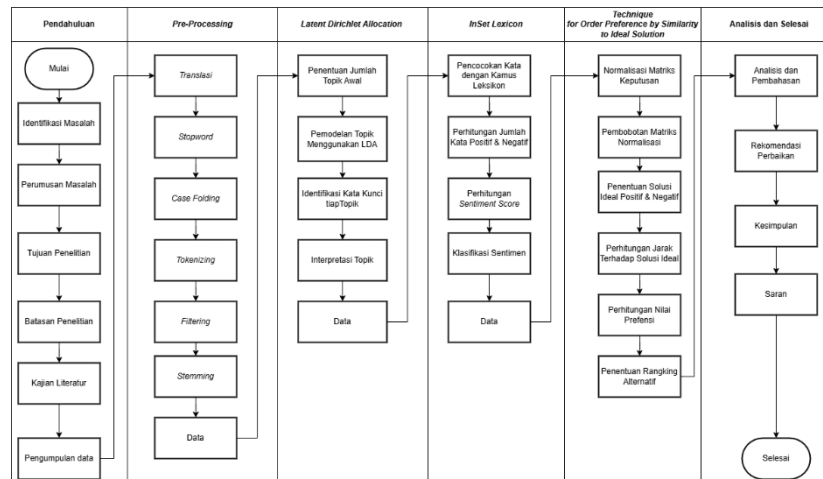
Objek penelitian ini adalah ulasan atau opini pengguna tentang destinasi wisata HeHa *Forest* yang dikumpulkan dari *platform* digital. HeHa *Forest* sendiri merupakan objek wisata alam yang memadukan konsep restoran, *playground*, *mini zoo*, hingga *Spot* foto yang *instagramable*. Ulasan yang dihasilkan dari *gap* antara ekspektasi dan realita ini mencerminkan persepsi, pengalaman, dan penilaian wisatawan tentang berbagai aspek layanan yang tersedia bagi mereka.

3.3 Jenis & Sumber Data

Data kualitatif yang digunakan dalam penelitian ini adalah teks ulasan wisatawan yang diolah menjadi data kuantitatif melalui proses analisis. Data sekunder ini diperoleh dari *platform Google Maps* dan *Instagram* dari tahun 2023 hingga Februari 2026, dan dikumpulkan melalui teknik *web scraping* menggunakan *Python*.

3.4 Alur Penelitian

Berikut merupakan alur penelitian yang diterapkan pada penelitian ini:



Gambar 3. 1 Alur penelitian

Alur penelitian dalam penelitian ini disusun secara sistematis untuk menggambarkan tahapan yang dilakukan mulai dari tahap awal hingga menghasilkan rekomendasi. Berdasarkan diagram alir penelitian, tahapan penelitian dijelaskan sebagai berikut:

1. Mulai
2. Identifikasi Masalah

Tahap ini dilakukan untuk mengidentifikasi masalah yang terjadi pada objek penelitian, yaitu belum adanya analisis yang terstruktur tentang persepsi wisatawan terhadap destinasi HeHa *Forest* berdasarkan data ulasan daring. Meskipun ulasan wisatawan dapat dianalisis secara efektif, mereka belum digunakan secara efektif dalam proses pengambilan keputusan berbasis data.

3. Perumusan Masalah

Selanjutnya, masalah yang telah ditemukan dirumuskan untuk menentukan fokus penelitian. Mengidentifikasi sentimen wisatawan, menemukan topik utama ulasan menggunakan metode LDA, dan menggunakan metode TOPSIS untuk menentukan prioritas perbaikan layanan adalah masalah yang dibahas dalam penelitian ini.

4. Tujuan Penelitian

Tujuan dari tahap ini adalah untuk menetapkan tujuan penelitian yang ingin dicapai. Tujuannya adalah untuk menganalisis sentimen wisatawan, menemukan topik utama yang dibahas dalam ulasan, dan secara kuantitatif menentukan aspek mana yang paling penting untuk perbaikan layanan.

5. Kajian Literatur

Kajian literatur dilakukan untuk mendukung penelitian dengan teori dan metode seperti *text mining*, analisis sentimen berbasis leksikon, *topic modelling*

menggunakan LDA, dan metode pengambilan keputusan TOPSIS. Tahap ini bertujuan untuk memberikan dasar ilmiah untuk metode yang digunakan dalam penelitian.

6. Batasan Penelitian

Tahap ini dilakukan untuk meningkatkan fokus penelitian. Penggunaan metode analisis sentimen berbasis *InSet Lexicon*, *topic modelling* menggunakan LDA, dan metode TOPSIS untuk menentukan prioritas perbaikan layanan termasuk penggunaan data ulasan dari *platform Instagram* dan *Google Maps* dari tahun 2023 hingga Februari 2026.

7. Pengumpulan Data

Metode *web scraping* digunakan untuk mengumpulkan data ulasan wisatawan dari *Instagram* dan *Google Maps*. Data yang dikumpulkan terdiri dari komentar, teks ulasan, dan deskripsi pengalaman wisatawan tentang destinasi *HeHa Forest*. Proses pengumpulan data dilakukan secara efisien menggunakan bahasa pemrograman *Python*. Setelah data dikumpulkan, mereka kemudian disimpan dalam bentuk kumpulan data yang akan digunakan dalam tahap analisis berikutnya.

8. Pre-processing Data

Proses *pre-processing* dilakukan untuk membersihkan dan menyiapkan data teks untuk analisis. Data ulasan yang diperoleh dari proses *scraping* masih tidak terstruktur dan mengandung suara, jadi perlu dilakukan beberapa tahapan berikut:

a. Translate

Translasi kata-kata dari Bahasa asal ke Bahasa Inggris lalu dikembalikan ke Bahasa Indonesia agar kata-kata tersebut menjadi dapat dipahami leksikon.

b. Stopword

Menghapus kata-kata yang tidak memiliki makna penting seperti “dan”, “yang”, “di”, dan lain-lain.

c. Case Folding

Mengubah seluruh huruf dalam teks menjadi huruf kecil untuk menyamakan format penulisan

d. Tokenizing

Memecah teks menjadi *unit* kata atau token sehingga dapat diproses lebih lanjut.

e. Filtering

Penghapusan emotikon, simbol, *mention* (@nickname).

f. *Stemming*

Mengubah kata menjadi bentuk dasarnya menggunakan *library* Sastrawi, misalnya “bermain” menjadi “main”.

9. *Topic Modelling (LDA)*

Metode ini digunakan untuk mengidentifikasi topik utama dalam ulasan wisatawan berdasarkan distribusi kata. Tahapan LDA meliputi:

a. Penentuan jumlah topik awal

Menentukan rentang jumlah topik yang akan diuji dalam proses pemodelan.

b. Perhitungan *coherence score*

Mengukur kualitas topik yang dihasilkan untuk setiap jumlah topik yang diuji.

c. Pemilihan jumlah topik optimal

Memilih jumlah dengan nilai *coherence* tertinggi dengan batas minimal 10 topik.

d. Pemodelan topik menggunakan LDA

Mengelompokkan kata ke dalam topik berdasarkan probabilitas kemunculan kata.

e. Identifikasi kata kunci tiap topik

Menentukan kata-kata yang merepresentasikan masing-masing topik.

f. Interpretasi topik

Memberikan makna pada setiap topik sesuai dengan konteks penelitian.

10. Interpretasi Topik

Kemudian, masalah yang dihasilkan dari metode LDA perlu pemberian makna pada tiap topik. Untuk melakukan ini, kata kunci dari setiap topik diinterpretasikan dan dihubungkan dengan kategori aspek yang sesuai. Tujuan dari langkah ini adalah untuk membuat hasil *topic modelling* lebih mudah untuk digunakan dalam analisis lanjutan yang lebih terorganisir. Hasil dari topik tersebut juga bertujuan sebagai alternatif TOPSIS

11. Analisis Sentimen (*InSet Lexicon*)

Analisis sentimen dilakukan setelah hasil topik diinterpretasi, sehingga nilai sentimen dapat dihitung berdasarkan masing-masing aspek yang telah diidentifikasi. Tahapan analisis sentimen adalah sebagai berikut:

a. Pencocokan kata dengan kamus leksikon

Setiap kata dalam ulasan dibandingkan dengan kamus *InSet Lexicon* untuk menentukan polaritasnya.

- b. Perhitungan jumlah kata positif dan negatif

Kata yang termasuk kategori positif dan negatif dihitung berdasarkan hasil pencocokan.

- c. Perhitungan *sentiment score*

$$Sentiment Score = \sum Kata Positif - \sum Kata Negatif \quad (3.1)$$

- d. Klasifikasi sentimen

Nilai yang dihasilkan digunakan untuk mengelompokkan sentimen menjadi positif, negatif, atau netral pada masing-masing aspek.

12. Perhitungan TOPSIS

Metode TOPSIS digunakan untuk menentukan prioritas aspek layanan yang perlu diperbaiki berdasarkan hasil analisis sebelumnya. Tahapan dalam metode TOPSIS sebagai berikut:

- a. Penyusunan matriks Keputusan

Matriks disusun berdasarkan alternatif yang berupa hasil interpretasi dan kriteria (*sentiment score* dan frekuensi topik).

- b. Normalisasi matriks Keputusan

Berikut merupakan rumus perhitungan normalisasi matriks keputusan sesuai dengan persamaan 2.1:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}^2} \quad (2.1)$$

- c. Pembobotan matriks normalisasi

Berikut merupakan rumus pada pembobotan matriks normalisasi sesuai dengan persamaan 2.2:

$$v_{ij} = w_{ij} \times r_{ij} \quad (2.2)$$

- d. Penentuan solusi ideal positif dan negatif

Solusi ideal positif diperoleh dari nilai maksimum, sedangkan solusi ideal negatif diperoleh dari nilai minimum untuk kriteria keuntungan.

- e. Perhitungan jarak terhadap solusi ideal

Berikut merupakan rumus jarak solusi ideal sesuai dengan persamaan 2.5 dan 2.6:

$$D_i^+ = \sqrt{\sum_{j=1}^n (y_i^+ - y_{ij})^2} \quad (2.5)$$

$$D_i^- = \sqrt{\sum_{j=1}^n (y_i^- - y_{ij})^2} \quad (2.6)$$

f. Perhitungan nilai prefensi

Berikut merupakan rumus nilai prefensi sesuai dengan persamaan 2.7:

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad (2.7)$$

g. Penentuan rangking alternatif

Alternatif dengan nilai preferensi tertinggi menjadi prioritas utama perbaikan.

13. Analisis dan Pembahasan

Pada tahap ini, Untuk menjawab rumusan masalah penelitian, analisis sentimen, LDA, dan *InSet Lexicon* diinterpretasikan. Selain itu, Hasil perhitungan yang diperoleh dianalisis untuk mengetahui prioritas aspek yang perlu diperbaiki.

14. Rekomendasi Perbaikan.

Berisi rekomendasi perbaikan yang terbentuk untuk meningkatkan kualitas pada wisata HeHa *Forest*.

15. Kesimpulan

Pada tahap akhir penelitian ini, tahap dimana memberikan jawaban dalam rumusan masalah berdasarkan hasil yang telah dilakukan.

16. Saran

Tahap saran ini, berisi rekomendasi-rekomendasi kepada peneliti selanjutnya.

17. Selesai

BAB IV

PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Pengumpulan Data

Data yang dikumpulkan untuk penelitian ini dikumpulkan menggunakan teknik *web scraping* untuk mengumpulkan ulasan pengunjung tentang destinasi wisata HeHa Forest. Data ini diambil dari *platform* digital seperti *Instagram* dan *Google Maps*, yang merupakan konten yang dibuat oleh pengguna.

4.1.1 Data Google Maps

Pengumpulan data dari *Google Maps* HeHa Forest dilakukan secara otomatis menggunakan metode *web scraping* dari interaksi dengan *Application Programming Interface* (API) pada *website Apify* dengan nama *Google Maps Review Scraper*. Metode ini dipilih untuk efisiensi dalam mengumpulkan *volume* data besar serta cepat. Berikut merupakan Langkah-langkah detail pengumpulan data dari *google maps*:

1. *Scraping* menggunakan *website apify Google Maps Review Scraper*

Pada Langkah pertama yaitu dengan membuka *website apify* lalu mencari judul *Google Maps Review Scraper* dengan *actor Compass*. Setelah ditemukan, *copy link maps* HeHa Forest dan *paste* ke *field* yang sudah disediakan serta tentukan berapa jumlah *review* yang ingin di *scraping*. Setelah dipastikan sesuai, klik *start* untuk menjalankan *apify* tersebut.

2. *Export* hasil

Hasil yang didapatkan dapat di *export* menjadi *file* yang kita inginkan, tahap ini pilih “*xlsx*” agar menjadi *file excel* agar dapat dilihat secara jelas. Berikut merupakan beberapa contoh hasil dari *apify*:

Tabel 4. 1 Data Review Google Maps

<i>title</i>	<i>url</i>	<i>name</i>	<i>text</i>
HeHa Fores t	https://www.google.com/maps/search/?api=1&query=HeHa%20Forest&query_place_id=ChIJ82xdbMVhei4RCIlgBIxnSW8	Blasius Febrianus	Restoran kurang variatif menunya, pesan americano dikasih es kopi pake gula.. Pizza pesen keras dan asin.. Semoga diperbaiki pelayanan restonya atau tutup saja restonya dan fokus ke camping
HeHa Fores t	https://www.google.com/maps/search/?api=1&query=HeHa%20Forest&query_place_id=ChIJ82xdbMVhei4RCIlgBIxnSW8	Me Min	Anakku korban Tiktok. Tempatnya not worth it banget. kecil, makanannya juga biasa aja dan harganya sangat mahal. Nggak ada specialnya. Yang enak di sini untukku hanya angin nya adem jadi betah duduk di lantai 2 nya tapi buat apa duduk aja kan? Kalau di tanya bakalan balik kah? Big No sih jawabannya
HeHa Fores t	https://www.google.com/maps/search/?api=1&query=HeHa%20Forest&query_place_id=ChIJ82xdbMVhei4RCIlgBIxnSW8	Annisa Krisridwan y	Biaya masuk free, untuk biaya Playground per anak 20 ribu, bersih, dan waktu kesana Playground masih sepi (hari libur, jam 11 siang). Ada mainan yang bisa dipakai juga. Ada beberapa <i>Spot</i> yang masih ditutup untuk perbaikan. Pesan menu di cafenya lumayan, pas ada promo juga. Bisa camping juga disana.
HeHa Fores t	https://www.google.com/maps/search/?api=1&query=HeHa%20Forest&query_place_id=ChIJ82xdbMVhei4RCIlgBIxnSW8	Galuh Puspitasari	Kesini htm free, masuk playground cuma bayar 20K udah free pendamping. Kemarin kesana jam 10an masih sepi, berasa private playground tp setelah itu trus rame Banyak <i>Spot</i> foto, kalo dateng sore/malem kayaknya lebih seru, bisa nginep di glamping juga

4.1.2 Data Instagram

Pengumpulan data dari *Instagram HeHa Forest* melalui proses *web scraping* komentar pada masing-masing *postingan Instagram HeHa Forest* yang mencerminkan opini serta pengalaman pengunjung. Kelemahan pada pengumpulan data menggunakan proses ini yaitu tidak ter-input nya balasan komentar pada *post Instagram HeHa Forest* sehingga yang dapat di *scrapping* hanya komentar utama. Berikut merupakan langkah-langkah dalam mengumpulkan data *Instagram*:

1. *Install Visual Studio Code dan Install Python*

Pada tahap pertama dalam mengumpulkan data *Instagram*, *visual studio code* dan *python* merupakan aplikasi tambahan agar dapat mengoptimalkan *scraper data comment Instagram*.

2. *Masukan code*

Dalam memasukkan *code* memerlukan *link website apify* yaitu *Instagram comments scraper* dengan *actor scrapesmith* agar lebih optimal dalam *scrapping*. Tahapan selanjutnya, *copy link per-postingan* pada *Instagram HeHa Forest* lalu cantumkan ke dalam *coding visual studio code*. Setelah dilakukan nya *code*, klik *run* agar dapat menjalankan *coding* yang akan mendapatkan hasilnya

3. *Export hasil*

Setelah didapatkan hasil tersebut, *export code* ke *format xlsx*. Contoh data yang dihasilkan dapat dilihat pada tabel dibawah:

Tabel 4. 2 Data *comment instagram*

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
	https://www.instagram.com/reel/DVSwGMEjpHW/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==		https://www.instagram.com/reel/DVSwGMEjpHW/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==					https://scontent.cdninstagram.com/v/t51.82787-19/621804637_18523054651065670_4537536154449431026_n.jpg?stp=dst-jpg_s150x150_tt6&_nc_cat=102&ccb=7-5&_nc_sid=f7ccc5&efg=eyJ2ZW5jb2RlX3RhZyI6InByb2ZpbGVfcGljLnd3dy4xMDgwLkMzIn0%3D&_nc_ohc=8XqMudh3sbUQ7kNvwEDSzPS&_nc_oc=Adl7YXCnjwtGqZzYljBJvWVG8k-lSXwSTfDpJmZcNOsoYo66QnZGknmgds63-hscObk&_nc_zt=24&_nc_ht=scontent.cdninstagram.com&_nc_			
3,842	zRlODBi		ZA==/c/1809				Vera	Gknmgds63-	vera.k		
34E+	NWFlZA	1,80928	28172124731	17722635	20025	vera.kusu	Kusuma	hscObk&_nc_zt=24&_nc_ht=sc	usuma		enak banget
18	==	E+16	78/	72	77669	maa	Dewi	ontent.cdninstagram.com&_nc_	a	vera.kusumaa	nih

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
								gid=- LPS61bA2ZiQDs4h1NOYgg&_n c_ss=8&oh=00_AfwCQlyhsFh9 yW4VsBFjA0ehE- FBbILVieDxCKQlxbBoVQ&oe= 69ADEF84 https://scontent.cdninstagram.co m/v/t51.2885- 19/330400839_75559439234830 0_4240063359528594889_n.jpg ?stp=dst- jpg_s150x150_tt6&_nc_cat=102 &ccb=7- 5&_nc_sid=f7ccc5&efg=eyJ2Z W5jb2RlX3RhZyI6InByb2ZpbGV fcGljLnd3dy4xMDgwLkMzIn0%			
3,842	https://www.instagram.com/reel/DVSwGMEjpHW/?utm_source=ig_web_c		https://www.instagram.com/reel/DVSwGMEjpHW/?utm_source=ig_web_c		44549						
34E+	&igsh=M	1,79172	72141781240	17722637	03616	hehaFore	HeHa	3D&_nc_ohc=scYhfifgr-	hehaF		
18	zRIODBi	E+16	24/	62	6	st	Forest	kQ7kNvwHtEhf2&_nc_oc=Adk7	orest	hehaForest	😊

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
	NWFIZA							5xE3xGPPH92qzhmZT7r4zDEv6			
	==							_iiiK7xs7233K3_4EHrncPJhk-			
								Dd7CB8NV6lsY&_nc_zt=24&_nc_ht=scontent.cdninstagram.com&_nc_ss=8&oh=00_AfxVSoTJxjam_EBqazl64keMUGcbBeMylyJxW4k33gehbA&oe=69ADDF62			
	https://www.instagram.com/reel/DVSwGwGMEjpHW/?utm_source=		https://www.instagram.com/reel/DVSwGMEjpHW/?utm_source=					https://scontent.cdninstagram.com/v/t51.82787-19/642457122_18107986756691092_3050187643470769273_n.jpg?stp=dst-jpg_s150x150_tt6&_nc_cat=103&ccb=7-			
3,842	ig_web_c		nk&igsh=Mz		20754			5&_nc_sid=f7ccc5&efg=eyJ2Z			
34E+	opy_link	1,81698	RIODBiNWFZA==/c/1816	17722667	29909	rayhannad	rayhan	W5jb2RLX3RhZyI6InByb2ZpbGVfcGljLnd3dy4xMDgwLkMzIn0%	rayhan		duh
18	&igsh=M	E+16		27	1	i	nalendra	fcGljLnd3dy4xMDgwLkMzIn0%	nadi	rayhannadi	ngilerrr 😊

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
	zRIODBi		97740543943					3D&_nc_ohc=_hPMHXznbVgQ			
	NWFIZA		30/					7kNvwHt1UfE&_nc_oc=AdnhF			
	==							M6LdCfOjj8pOpfNuFgJVegASpl			
								pBqz6zu80yHAD CZ0gkGJSsksQ			
								HHWLvWAUA-			
								E&_nc_zt=24&_nc_ht=scontent			
								.cdninstagram.com&_nc_gid=-			
								LPS61bA2ZiQDs4h1NOYgg&_n			
								c_ss=8&oh=00_Afzh0iCds6VPu			
								oo807WU0rjdxih4VTqElEATQ1			
								ULVwIYHg&oe=69ADE00E			
	https://w		https://www.i				Kolabor	https://scontent.cdninstagram.co			Hai seluruh
	ww.instagram		nstagram.com				asi Jogja	m/v/t51.82787-			warga Jogja,
	ram.com/		/reel/DVSwG				Media	19/632086623_18018644258815			Kolaborasik
	reel/DVS		MEjpHW/?ut				Promosi	778_6967456642098702323_n.j			an Postingan
3,842	wGMEjp		m_source=ig		54284		&	pg?stp=dst-	kolabo		BERITA
34E+	HW/?utm	1,80691	_web_copy_li	17725047	43977	kolaborasi	Informa	jpg_sl50x150_tt6&_nc_cat=106	rasijog	kolaborasijogj	UPDATE,
18	_source=	E+16	nk&igsh=Mz	24	7	jogja	si	&ccb=7-	ja	a	PROMO

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
	<i>ig_web_c</i>		<i>RIODBiNWF1</i>				Yogyakarta	<i>5&_nc_sid=f7ccc5&efg=eyJ2Z</i>			PRODUK,
	<i>opy_link</i>		<i>ZA==/c/1806</i>				arta	<i>W5jb2RlX3RhZyI6InByb2ZpbGV</i>			PERSONAL
	<i>&igsh=M</i>		<i>90727012366</i>					<i>fcGljLnd3dy4xMDgwLkMzIn0%</i>			BRANDIN
	<i>zRIODBi</i>		<i>01/</i>					<i>3D&_nc_ohc=Rz1gCMUmBcoQ</i>			G, DLL
	<i>NWF1ZA</i>							<i>7kNvwFkJB1U&_nc_oc=AdkA</i>			anda di
	<i>==</i>							<i>MtakZmgXaITFGIwZYwOVtzMy</i>			@kolaborasi
								<i>egklEqF8lzZvryR2yKx99kL18yu6</i>			jogja
								<i>q0N7c1p4Vhk&_nc_zt=24&_nc</i>			GRATISSS!
								<i>_ht=scontent.cdninstagram.com</i>			
								<i>&_nc_gid=-</i>			
								<i>LPS61bA2ZiQDs4h1NOYgg&_n</i>			
								<i>c_ss=8&oh=00_AfzEULty2hunL</i>			
								<i>u9uXnipATbaqU-</i>			
								<i>FY1JfvpAVXukfRcdR7Q&oe=69</i>			
								<i>AE0468</i>			
3,841	<i>https://w</i>		<i>https://www.i</i>					<i>https://scontent.cdninstagram.co</i>			
99E+	<i>ww.instagram</i>	1,8071E	<i>nstagram.com</i>	17722406	22758			<i>m/v/t51.89012-</i>	dwlees		wahh enak
18	<i>ram.com/</i>	+16	<i>/reel/DVRefZ</i>	29	94381	dwleestari	dwi	<i>19/573323465_12198254633022</i>	tari	dwleestari	banget

<i>postId</i>	<i>postUrl</i>	<i>commentId</i>	<i>commentUrl</i>	<i>timestamp</i>	<i>userId</i>	<i>username</i>	<i>userFullName</i>	<i>ownerProfilePicUrl</i>	<i>owner</i>	<i>ownerUsername</i>	<i>text</i>
	reel/DVR		FkyaL/?utm_s					12_7278921664109726296_n.jp			pengen
	efZFkyaL		ource=ig_we					g?stp=dst-			kesana
	/?utm_so		b_copy_link&					jpg_s150x150_tt6&_nc_cat=1&			
	urce=ig_		igsh=MzRlO					ccb=7-			
	web_copy		DBiNWFlZA					5&_nc_sid=f7ccc5&efg=eyJ2Z			
	_link&igs		==/c/180709					W5jb2RlX3RhZyI6InByb2ZpbGV			
	h=MzRlO		64549193539/					fcGljLnd3dy5DMYJ9&_nc_ohc=			
	DBiNWFl							SMXWddwmmIcQ7kNvwH_kVtk			
	ZA==							&_nc_oc=Adn8WDfPYQUCzE_			
								htDy3eI5GBZCmvj2ERk7pDAS			
								QTKO0fiVHJhu-			
								dcw6FM2z4bxxvbQ&_nc_zt=24			
								&_nc_ht=scontent.cdninstagram			
								.com&_nc_gid=r-			
								ASITQOanIENdOEKVSJ8Q&_			
								nc_ss=8&oh=00_Afyd0nM7uGB			
								DHpY1FylgjIEcTOHGgbJ9SrjA			
								IUXWF0dROw&oe=69ADE422			

4.2 Pre-processing Data

Pre-processing data merupakan tahapan yang perlu diperhatikan dalam menganalisis teks untuk dibersihkan, menormalisasi, dan mengubah data mentah menjadi format yang sesuai agar dapat dianalisis lebih lanjut.

4.2.1 Translasi

Proses translasi ini digunakan agar dapat menghasilkan kata yang dapat dipahami leksikon. Translasi ini melalui 2 tahapan dari Bahasa asal menjadi Bahasa Inggris lalu dikembalikan lagi ke Bahasa Indonesia agar optimal. Berikut merupakan contoh proses translasi:

Tabel 4. 3 Translasi

Bahasa asal	Bahasa Inggris	Bahasa Indonesia
Wah harga ramah kantongggg	<i>Wow, the price is pocket-friendly</i>	Wow, harganya sangat terjangkau.

4.2.2 Stopword

Pada tahapan *Stopword* merupakan kata-kata umum yang sering muncul dalam suatu Bahasa namun memiliki informasi yang rendah, proses penghapusan *stopword* memiliki tujuan agar dapat mengurangi kompleksitas data serta memusatkan pada kata-kata penting yang lebih relevan. Berikut merupakan contoh penghapusan data tersebut;

Tabel 4. 4 Data *stopword*

Data sebelum	Data Setelah <i>stopword</i>
Wow, harganya sangat terjangkau.	Wow, harganya terjangkau.

4.2.3 Pengubahan huruf kecil (*Case Folding*)

Pada tahap ini, pengubahan semua huruf dalam teks diubah ke huruf kecil semua dengan tujuan agar menyeragamkan representasi kita:

Tabel 4. 5 Data *case folding*

Data sebelum	Data Setelah <i>case folding</i>
Wow, harganya terjangkau.	wow, harganya terjangkau.

4.2.4 Tokenisasi

Tahap tokenisasi merupakan pemisahan antar kata dalam satu teks yang disebut token. Dalam konteks ini, pemisahan kata tersebut dianggap sebagai satu token tersendiri. Berikut merupakan contoh pada tokenisasi:

Tabel 4. 6 Data tokenisasi

Data sebelum	Data setelah tokenisasi
wow, harganya terjangkau.	['wow', ',', 'harganya', 'terjangkau', '.']

4.2.5 Penghapusan data subjektif, karakter non-alphabet (filtering)

Pada tahapan ini bertujuan untuk penghapusan elemen-elemen yang bersifat subjektif atau hal-hal yang tidak memiliki relevansi untuk analisis sentimen. Penghapusan karakter *non alphabet* juga diperlukan karena tidak berkontribusi dalam makna leksikal. Data tersebut meliputi penghapusan emotikon, *username mention* (@username), tanda baca, angka, dan simbol karena tidak berkontribusi pada inti makna teks. *Filtering* juga bisa digunakan untuk menghilangkan kata-kata yang subjektif. Berikut merupakan hasil dari *filtering*:

Tabel 4. 7 Data Filtering

Data sebelum	Data setelah filtering
['wow', ',', 'harganya', 'terjangkau', '.']	['harganya', 'terjangkau']

4.2.6 Stemming

Stemming merupakan proses mengubah sebuah kata ke bentuk dasarnya. Sebagai contoh pada data dibawah ini, kata “harganya” menjadi “harga”. Tujuan dari tahap ini adalah untuk menyatukan berbagai turunan kata yang memiliki dasar yang sama, sehingga dapat dianggap sebagai satu entitas dalam analisis. Berikut merupakan contoh *stemmed*:

Tabel 4. 8 Data Stemmed

Data sebelum	Data setelah Stemmed
['harganya', 'terjangkau']	['harga']

4.2.7 Penyusunan kalimat

Setelah melewati tahapan-tahapan pada *pre-processing*, token-token tersebut disatukan Kembali menjadi sebuah kalimat. Langkah pada tahapan ini bertujuan untuk mempermudah proses lanjutan seperti *sentiment analysis*, yang memerlukan masukan berupa kalimat yang sudah ternormalisasi.

Tabel 4. 9 Data Penyusunan Kalimat

Data Stemmed	Kalimat Normalisasi
['harga']	harga

4.3 Pemodelan Topik (*Latent Dirichlet Allocation/LDA*)

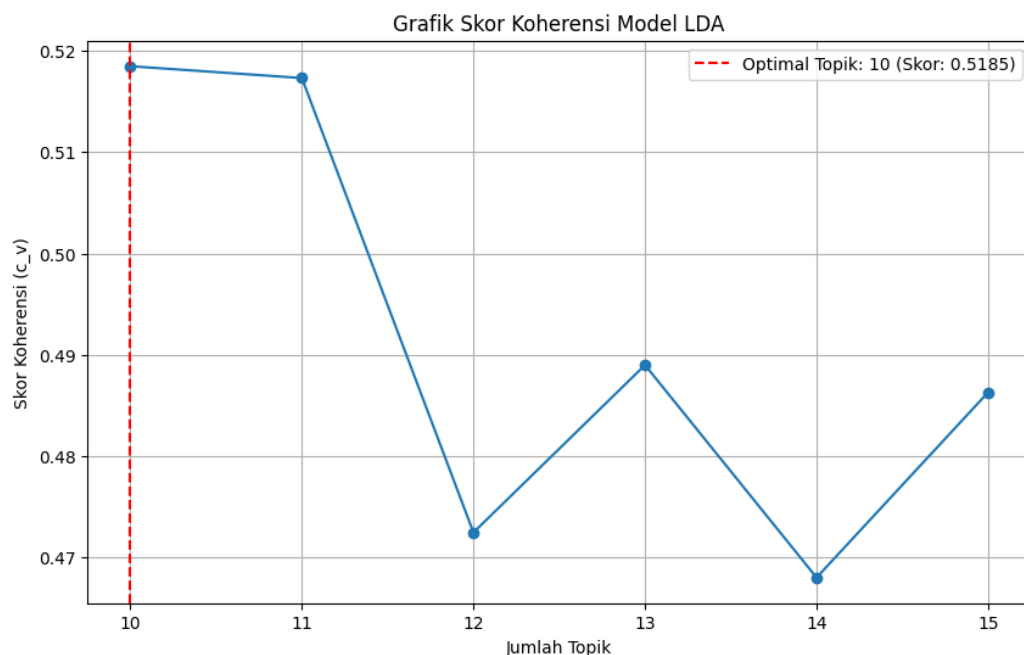
Dalam menemukan tema utama yang muncul pada ulasan pengunjung HeHa Forest, metode *Latent Dirichlet Allocation* (LDA) digunakan untuk pemodelan topik. Data yang

digunakan berasal dari hasil *pre-processing* dari ulasan di *Instagram* dan *Google Maps*. Proses pemodelan dilakukan dengan menguji berbagai topik dan kemudian memilih jumlah topik yang paling cocok berdasarkan nilai *coherence*. Model dengan nilai *coherence* tertinggi dipilih sebagai jumlah topik yang optimal, sehingga nilai *coherence* pada *google maps* berjumlah 0.5185 terpilih dengan jumlah 10 topik dan nilai *coherence* pada *Instagram* berjumlah 0.5906 juga terpilih dengan jumlah 11 topik.

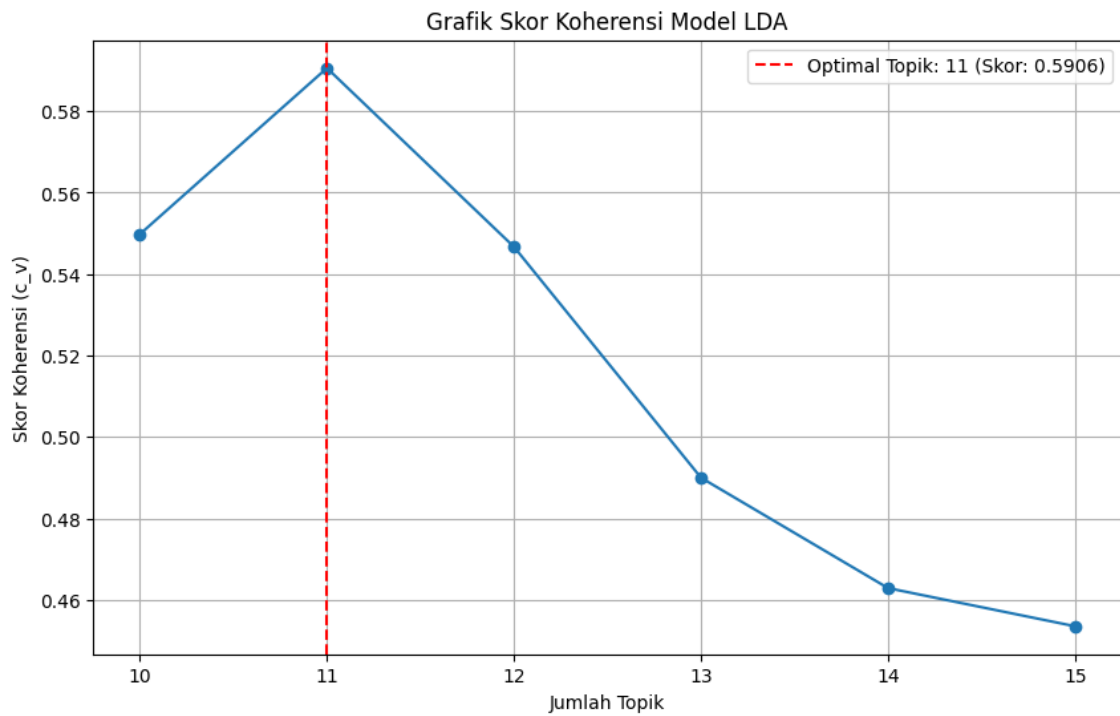
4.3.1 Hasil pemodelan topik

Hasil pemodelan topik menggunakan metode LDA menghasilkan sejumlah topik yang mewakili pola pembahasan dalam data ulasan. Untuk mengetahui bagaimana karakteristik ulasan berbeda pada masing-masing *platform*, proses pemodelan ini dilakukan secara terpisah pada data *Google Maps* dan *Instagram*.

Menguji berbagai variasi topik dan memilih model dengan nilai *coherence* terbaik digunakan untuk menentukan jumlah topik yang tepat. Hasil pengujian menunjukkan bahwa jumlah topik yang ideal untuk data *Google Maps* merupakan 10 topik dan *Instagram* berjumlah 11 topik yang dapat dilihat pada grafik koherensi. Berikut merupakan grafik *coherence* dari masing-masing *platform*:



Gambar 4. 1 Grafik *coherence* Google Maps



Gambar 4. 2 Grafik *coherence Instagram*

Setiap topik diwakili oleh kumpulan kata dengan probabilitas tertinggi yang menunjukkan hubungan antar kata. Berikut merupakan tabel yang menunjukkan hasil pemodelan topik untuk masing-masing *platform*:

Tabel 4. 10 Topik *Google Maps*

<i>Topic</i>	<i>Top Words (Probability)</i>									
<i>Topic 0</i>	tarik (0.249)	anak (0.163)	toilet (0.131)	jawab (0.012)	kosong (0.012)	modern (0.008)	si (0.008)	golong (0.008)	gerak (0.008)	iga (0.007)
<i>Topic 1</i>	keluarga (0.340)	ramah (0.243)	staf (0.060)	segar (0.044)	rapi (0.026)	rekreasi (0.026)	hidang (0.017)	piknik (0.013)	gunung (0.012)	sibuk (0.008)
<i>Topic 2</i>	pokok (0.068)	sesat (0.029)	temonjo (0.017)	ben (0.017)	dolano (0.017)	lur (0.017)	do (0.017)	healingmu (0.017)	sederhana (0.016)	olahraga (0.012)
<i>Topic 3</i>	foto (0.374)	hutan (0.207)	<i>Spot</i> (0.160)	kafe (0.038)	utara (0.017)	bayang (0.009)	judul (0.007)	market (0.005)	sleman (0.004)	segitu (0.004)
<i>Topic 4</i>	nyaman (0.634)	kota (0.116)	eksotis (0.008)	hangat (0.006)	gemericik (0.006)	istirahat (0.005)	luang (0.005)	rata (0.005)	melayanani (0.005)	mantullll (0.003)
<i>Topic 5</i>	makan (0.454)	enak (0.168)	minum (0.077)	mobil (0.037)	musik (0.022)	bakso (0.011)	sudut (0.006)	porsi (0.006)	hampir (0.005)	suruh (0.005)
<i>Topic 6</i>	main (0.222)	taman (0.207)	motor (0.055)	relaksasi (0.035)	alpukat (0.013)	awat (0.013)	penasaran (0.012)	larang (0.011)	jus (0.011)	9 (0.010)

<i>Topic</i>	<i>Top Words (Probability)</i>									
<i>Topic 7</i>	lokasi (0.290)	nongkrong (0.157)	kendara (0.026)	populer (0.022)	kerabat (0.015)	bareng (0.015)	teh (0.015)	angkut (0.014)	bolakbalik (0.013)	rapat (0.010)
<i>Topic 8</i>	parkir (0.082)	restoran (0.082)	bayar (0.077)	ribu (0.070)	harga (0.064)	tiket (0.051)	mahal (0.040)	jemput (0.023)	biaya (0.022)	menu (0.019)
<i>Topic 9</i>	wisata (0.280)	jogja (0.242)	fasilitas (0.059)	rupiah (0.026)	normal (0.016)	langgan (0.012)	segi (0.012)	inti (0.011)	jl (0.010)	butuh (0.010)

Tabel 4. 11 *Topic 10 Instagram*

<i>Topic</i>	<i>Top Words (Probability)</i>									
<i>Topic 0</i>	jogja (0.224)	alami (0.046)	yogyakart a (0.038)	sulit (0.032)	tuhan (0.018)	catat (0.015)	rizt43 (0.015)	wisatawa n (0.009)	astaga (0.009)	gila (0.007)
<i>Topic 1</i>	hutan (0.198)	tiket (0.086)	shyfa1709 (0.023)	ralfhyfhas a (0.023)	antusias (0.022)	nabilaarditaaza hra (0.019)	azahranabilaar dita (0.019)	obat (0.017)	dingin (0.016)	info (0.016)
<i>Topic 2</i>	main (0.075)	musik (0.036)	gali (0.029)	jembatan (0.022)	set (0.021)	lagu (0.021)	pesta (0.019)	pusing (0.018)	taman (0.017)	otak (0.015)
<i>Topic 3</i>	minum (0.055)	air (0.027)	kopi (0.026)	balas (0.025)	mati (0.024)	terjun (0.021)	pesan (0.016)	kondisi (0.014)	kangen (0.014)	teh (0.013)

<i>Topic</i>		<i>Top Words (Probability)</i>								
<i>Topic 4</i>	lokasi (0.067)	inovasi (0.042)	hindar (0.039)	positif (0.037)	gagal (0.028)	ketidaksempur naan (0.027)	trauma (0.027)	basah (0.022)	nomor (0.015)	come (0.014)
<i>Topic 5</i>	harga (0.155)	ribu (0.031)	pasar (0.029)	minn (0.027)	lancar (0.026)	paket (0.024)	promo (0.021)	like (0.014)	tualang (0.013)	keluh (0.012)
<i>Topic 6</i>	foto (0.160)	tarik (0.068)	<i>Spot</i> (0.047)	hibur (0.026)	jenak (0.023)	murah (0.019)	penasaran (0.017)	butuh (0.017)	kencan (0.016)	tulis (0.014)
<i>Topic 7</i>	bayar (0.107)	keluarga (0.058)	segar (0.054)	aktivitas (0.045)	dm (0.018)	nasi (0.018)	kenang (0.014)	ruang (0.014)	sudut (0.014)	takjub (0.014)
<i>Topic 8</i>	gantung (0.030)	negatif (0.027)	follow (0.023)	normal (0.021)	terkadang (0.020)	wujud (0.020)	alir (0.020)	suap (0.020)	kanya (0.020)	nur (0.020)
<i>Topic 9</i>	waras (0.037)	haha (0.029)	ra (0.021)	anjing (0.021)	partisipasi (0.018)	waspada (0.018)	dazartgallery (0.015)	peapeeee (0.015)	noritayukeca pri (0.015)	nezzlie (0.015)
<i>Topic 10</i>	makan (0.129)	biaya (0.066)	enak (0.053)	anak (0.051)	wisata (0.047)	kota (0.030)	nyaman (0.024)	kampung (0.020)	menu (0.018)	muda (0.013)

4.4 Interpretasi Topik

Setiap topik yang dihasilkan oleh metode *Latent Dirichlet Allocation* (LDA) memiliki maknanya sendiri. Untuk mencapai tujuan ini, tahap interpretasi topik dilakukan. Proses interpretasi dilakukan secara manual dengan kata-kata dominan yang memiliki kemungkinan tertinggi pada masing-masing topik dan kemudian menghubungkannya dengan konteks ulasan pengunjung HeHa Forest. Dengan demikian, setiap topik tidak lagi hanya kumpulan kata, tetapi telah mewakili konteks atau pembahasan tertentu yang terkait dengan pengalaman pengunjung. Hasil interpretasi topik dalam penelitian ini digunakan sebagai dasar untuk proses analisis sentimen dan penentuan prioritas menggunakan metode TOPSIS. Dengan demikian, setiap topik yang diinterpretasikan dianggap sebagai topik alternatif yang akan dianalisis.

4.4.1 Interpretasi topik pada data Google Maps

Berikut merupakan hasil interpretasi aspek pada data *Google Maps*

Tabel 4. 12 Interpretasi topik pada data *Google Maps*

<i>Topic</i>	<i>Top Words (Probability)</i>										Interpretasi Topik
	tarik	anak	toilet	jawab	kosong	modern	si	golong	gerak	iga	Fasilitas
<i>Topic 0</i>	(0.249)	(0.163)	(0.131)	(0.012)	(0.012)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	bermain anak
	keluarga	ramah	staf	segar	rapi	rekreasi	hidang	piknik	gunung	sibuk	Keramahan
<i>Topic 1</i>	(0.340)	(0.243)	(0.060)	(0.044)	(0.026)	(0.026)	(0.017)	(0.013)	(0.012)	(0.008)	Staf
	pokok	sesat	temonjo	ben	dolano	lur	do	healingm	sederhana	olahraga	
<i>Topic 2</i>	(0.068)	(0.029)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	u (0.017)	(0.016)	(0.012)	Relaksasi

<i>Topic</i>	<i>Top Words (Probability)</i>										Interpretasi Topik
<i>Topic 3</i>	foto (0.374)	hutan (0.207)	Spot (0.160)	kafe (0.038)	utara (0.017)	bayang (0.009)	judul (0.007)	market (0.005)	sleman (0.004)	segitu (0.004)	Spot Foto Menarik
<i>Topic 4</i>	nyaman (0.634)	kota (0.116)	eksotis (0.008)	hangat (0.006)	gemericik (0.006)	istirahat (0.005)	luang (0.005)	rata (0.005)	melayanan i (0.005)	mantullll (0.003)	Kenyamanan
<i>Topic 5</i>	makan (0.454)	enak (0.168)	minum (0.077)	mobil (0.037)	musik (0.022)	bakso (0.011)	sudut (0.006)	porsi (0.006)	hampir (0.005)	suruh (0.005)	Makanan Enak
<i>Topic 6</i>	main (0.222)	taman (0.207)	motor (0.055)	relaksasi (0.035)	alpukat (0.013)	awat (0.013)	penasaran (0.012)	larang (0.011)	jus (0.011)	9 (0.010)	taman bermain
<i>Topic 7</i>	lokasi (0.290)	nongkron g (0.157)	kendara (0.026)	populer (0.022)	kerabat (0.015)	bareng (0.015)	teh (0.015)	angkut (0.014)	bolakbalik (0.013)	rapat (0.010)	Tempat Nongkrong
<i>Topic 8</i>	parkir (0.082)	restoran (0.082)	bayar (0.077)	ribu (0.070)	harga (0.064)	tiket (0.051)	mahal (0.040)	jemput (0.023)	biaya (0.022)	menu (0.019)	Biaya Parkir dan Makanan
<i>Topic 9</i>	wisata (0.280)	jogja (0.242)	fasilitas (0.059)	rupiah (0.026)	normal (0.016)	langgan (0.012)	segi (0.012)	inti (0.011)		butuh (0.010)	Fasilitas Umum

4.4.2 Interpretasi topik pada data Instagram

Berikut merupakan hasil interpretasi aspek pada data *Instagram*:

Tabel 4. 13 Interpretasi topik pada *Instagram*

<i>Topic</i>	<i>Top Words (Probability)</i>										Interpretasi Topik
<i>Topic 0</i>	jogja (0.224)	alami (0.046)	yogyak arta (0.038)	sulit (0.032)	tuhan (0.018)	catat (0.015)	rizt43 (0.015)	wisataw an (0.009)	astaga (0.009)	gila (0.007)	wisata dengan suasana alam
<i>Topic 1</i>	hutan (0.198)	tiket (0.086)	shyfa170 9 (0.023)	ralfhyfha sa (0.023)	antusias (0.022)	nabilaardita zahra (0.019)	azahranabila ardita (0.019)	obat (0.017)	dingin (0.016)	info (0.016)	suasana tempat (hutan)
<i>Topic 2</i>	main (0.075)	musik (0.036)	gali (0.029)	jembatan (0.022)	set (0.021)	lagu (0.021)	pesta (0.019)	pusing (0.018)	taman (0.017)	otak (0.015)	hiburan musik
<i>Topic 3</i>	minum (0.055)	air (0.027)	kopi (0.026)	balas (0.025)	mati (0.024)	terjun (0.021)	pesan (0.016)	kondisi (0.014)	kangen (0.014)	teh (0.013)	minuman
<i>Topic 4</i>	lokasi (0.067)	inovasi (0.042)	hindar (0.039)	positif (0.037)	gagal (0.028)	ketidaksemp urnaan (0.027)	trauma (0.027)	basah (0.022)	nomor (0.015)	come (0.014)	<i>improvement</i> lokasi
<i>Topic 5</i>	harga (0.155)	ribu (0.031)	pasar (0.029)	minn (0.027)	lancar (0.026)	paket (0.024)	promo (0.021)	like (0.014)	tualang (0.013)	keluh (0.012)	Harga dan promo
<i>Topic 6</i>	foto (0.160)	tarik (0.068)	<i>Spot</i> (0.047)	hibur (0.026)	jenak (0.023)	murah (0.019)	penasaran (0.017)	butuh (0.017)	kencan (0.016)	tulis (0.014)	<i>Spot</i> Foto menarik
<i>Topic 7</i>	bayar (0.107)	keluarg	segar (0.054)	aktivitas (0.045)	dm (0.018)	nasi (0.018)	kenang (0.014)	ruang (0.014)	sudut (0.014)	takjub (0.014)	Biaya Aktifitas Keluarga

<i>Topic</i>	<i>Top Words (Probability)</i>										Interpretasi Topik
	a (0.058)										
<i>Topic 8</i>	gantun g (0.030)	negatif (0.027)	follow (0.023)	normal (0.021)	terkadang g (0.020)	wujud (0.020)	alir (0.020)	suap (0.020)	kanya (0.020)	nur (0.020)	<i>Healing</i> dan Interaksi Pengguna
<i>Topic 9</i>	waras (0.037)	haha (0.029)	ra (0.021)	anjing (0.021)	partisipa si (0.018)	waspada (0.018)	dazartgaller y (0.015)	peapeeee (0.015)	noritayuke capri (0.015)	nezzlie (0.015)	Ekspresi Pengguna
<i>Topic 10</i>	makan (0.129)	biaya (0.066)	enak (0.053)	anak (0.051)	wisata (0.047)	kota (0.030)	nyaman (0.024)	kampung g (0.020)	menu (0.018)	muda (0.013)	menu makanan

4.5 Analisis Sentimen menggunakan *InSet Lexicon*

Pada Langkah ini perlu melakukan analisis sentimen terhadap ulasan pengunjung setelah mengidentifikasi elemen utama berdasarkan hasil pemodelan topik. Tujuan dari analisis ini adalah untuk mengetahui kecenderungan pendapat pengunjung tentang masing-masing elemen yang telah diidentifikasi. Untuk membedakan sentimen menjadi positif, negatif, dan netral, pendekatan berbasis *InSet Lexicon* digunakan.

4.5.1 Metode analisis sentiment

Analisis sentimen dilakukan untuk mengetahui preferensi pengunjung terhadap elemen yang telah diperoleh dari pemodelan topik. Metode berbasis leksikon digunakan dengan *InSet Lexicon*, kamus sentimen berbahasa Indonesia yang mencakup daftar kata dan bobot polaritasnya. Setiap ulasan yang telah melewati tahap *pre-processing* dievaluasi dengan membandingkan kata-kata yang dihasilkan stemming dengan kamus *InSet Lexicon*. Setiap ulasan memiliki bobot kata positif dan negatif. Nilai sentimen kemudian dimasukkan ke dalam tiga kategori: positif, negatif, dan netral.

4.5.2 Hasil analisis sentiment Instagram

Hasil analisis sentimen disusun berdasarkan topik yang telah ditentukan pada tahap sebelumnya. Berikut merupakan jumlah rekapitulasi *sentiment Instagram*:

Tabel 4. 14 Rekapitulasi sentimen *Instagram*

<i>Topic</i>	<i>Negative</i>	<i>Neutral</i>	<i>Positive</i>	<i>Rata-rata sentiment per topic</i>
<i>Topic 0</i>	7	282	18	0,03583
<i>Topic 1</i>	3	73	9	0,07059
<i>Topic 2</i>	0	75	3	0,03846
<i>Topic 3</i>	1	9	0	-0,1
<i>Topic 4</i>	1	27	0	-0,0357
<i>Topic 5</i>	1	82	0	-0,012
<i>Topic 6</i>	0	86	5	0,05495
<i>Topic 7</i>	5	72	0	-0,0649
<i>Topic 8</i>	0	5	1	0,16667
<i>Topic 9</i>	0	2	0	0
<i>Topic 10</i>	2	61	6	0,05797

Hasil rekapitulasi menunjukkan bahwa sentimen yang didistribusikan di *Instagram* menunjukkan perbedaan pendapat di setiap topik. Topik 1 memiliki nilai tertinggi yaitu 0,07059 dan topik 9 berada diposisi terakhir yaitu 0.

4.5.3 Hasil analisis sentiment Google Maps

Hasil analisis sentimen disusun berdasarkan topik yang telah ditentukan pada tahap sebelumnya. Berikut merupakan jumlah rekapitulasi *sentiment Google Maps*:

Tabel 4. 15 Rekapitulasi sentimen *Google Maps*

<i>Topic</i>	<i>Negative</i>	<i>Neutral</i>	<i>Positive</i>	<i>Rata-rata sentiment per topic</i>
<i>Topic 0</i>	2	80	31	0,336
<i>Topic 1</i>	6	163	140	0,566
<i>Topic 2</i>	2	5	5	0,25
<i>Topic 3</i>	21	212	113	0,358
<i>Topic 4</i>	1	4	441	1,302
<i>Topic 5</i>	24	99	187	0,638
<i>Topic 6</i>	5	40	8	0,075
<i>Topic 7</i>	3	19	3	0
<i>Topic 8</i>	3	13	4	0,05
<i>Topic 9</i>	1	40	2	0,02

Hasil rekapitulasi menunjukkan bahwa sentimen netral mendominasi setiap topik *Google Maps*. Topik 4 memiliki nilai tertinggi yaitu 1,302 dan topik 7 berada diposisi terakhir yaitu 0.

4.6 TOPSIS

4.6.1 Penentuan alternatif dan kriteria

Dalam penelitian ini, penentuan prioritas topik perlu diperhatikan berdasarkan hasil analisis sentimen. Alternatif yang digunakan adalah topik hasil interpretasi dan Kriteria yang digunakan dalam perhitungan TOPSIS, terdiri dari nilai rata-rata sentimen dan jumlah ulasan negatif. Nilai rata-rata sentimen menunjukkan tingkat persepsi pengunjung terhadap setiap elemen, sedangkan jumlah ulasan negatif menunjukkan tingkat masalah yang muncul pada elemen tersebut. Tujuan utama dalam penelitian ini adalah menentukan prioritas perbaikan aspek, sehingga jenis kriteria ditentukan sebagai berikut:

1. Jumlah sentimen negatif sebagai kriteria benefit, karena semakin tinggi nilai negatif maka menunjukkan aspek tersebut memiliki banyak keluhan yang muncul sehingga semakin perlu diperbaiki.

2. Nilai rata-rata sentimen sebagai kriteria *cost*, karena semakin rendah nilai sentimennya maka persepsi pengunjung semakin buruk sehingga semakin perlu diperbaiki karena memerlukan perhatian lebih besar.

4.6.2 Pembentukan matriks Keputusan

Hasil analisis sentimen pada setiap topik membentuk matriks keputusan yang digunakan sebagai input dalam metode TOPSIS. Matriks keputusan tersebut terdiri atas dua kriteria, yaitu jumlah komentar sentimen negatif dan nilai sentimen rata-rata pada setiap topik. Jumlah komentar sentimen negatif digunakan untuk merepresentasikan tingkat keluhan pengunjung terhadap suatu topik, sedangkan nilai sentimen rata-rata digunakan untuk menggambarkan kecenderungan persepsi pengunjung secara keseluruhan melalui agregasi skor sentimen pada setiap topik. Penggunaan nilai sentimen rata-rata memungkinkan proses pengambilan keputusan tidak hanya mempertimbangkan intensitas keluhan, tetapi juga mempertimbangkan kecenderungan opini pengunjung secara umum sehingga menghasilkan penilaian yang lebih representatif terhadap kondisi setiap aspek (Birjali dkk., 2021). Berdasarkan kedua kriteria tersebut, diperoleh matriks keputusan yang selanjutnya digunakan sebagai dasar perhitungan menggunakan metode TOPSIS. Berikut merupakan tabel matriks keputusan *Google Maps* dan *Instagram*:

Tabel 4. 16 Matriks Keputusan *Google Maps*

<i>Topic</i>	<i>Negative</i>	<i>Rata-rata sentiment per topic</i>
<i>Topic 0</i>	2	0,336
<i>Topic 1</i>	6	0,566
<i>Topic 2</i>	2	0,25
<i>Topic 3</i>	21	0,358
<i>Topic 4</i>	1	1,302
<i>Topic 5</i>	24	0,638
<i>Topic 6</i>	5	0,075
<i>Topic 7</i>	3	0
<i>Topic 8</i>	3	0,05
<i>Topic 9</i>	1	0,02

Tabel 4. 17 Matriks Keputusan *Instagram*

<i>Topic</i>	<i>Negative</i>	<i>Rata-rata sentiment per topik</i>
<i>Topic 0</i>	7	0,03583
<i>Topic 1</i>	3	0,07059
<i>Topic 2</i>	0	0,03846
<i>Topic 3</i>	1	-0,1
<i>Topic 4</i>	1	-0,0357
<i>Topic 5</i>	1	-0,012
<i>Topic 6</i>	0	0,05495
<i>Topic 7</i>	5	-0,0649
<i>Topic 8</i>	0	0,16667
<i>Topic 9</i>	0	0
<i>Topic 10</i>	2	0,05797

4.6.3 Normalisasi dan pembobotan

Normalisasi terhadap matriks Keputusan digunakan untuk menghilangkan perbedaan skala antar kriteria. *Sentiment* negatif diartikan sebagai *benefit* yang berarti semakin tinggi maka semakin perlu diperbaiki. Rata-rata *sentiment*/skor diartikan sebagai *cost* yang berarti semakin rendah maka semakin perlu diperbaiki. Berikut merupakan rumus perhitungan normalisasi matriks Keputusan pada persamaan 2.1:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}^2} \quad (2.1)$$

Adapun normalisasi *Google Maps* adalah sebagai berikut:

$$negatif = \sqrt{2^2 + 6^2 + 2^2 + 21^2 + 1^2 + 24^2 + 5^2 + 3^2 + 3^2 + 1^2} = 33,25$$

$$skor = \sqrt{0,336^2 + 0,566^2 + 0,25^2 + 0,358^2 + 1,302^2 + 0,638^2 + 0,075^2 + 0^2 + 0,05^2 + 0,02^2} = 1,65$$

$$\begin{aligned} negatif &= \left(\frac{2}{33,25}\right), \left(\frac{6}{33,25}\right), \left(\frac{2}{33,25}\right), \left(\frac{21}{33,25}\right), \left(\frac{1}{33,25}\right), \left(\frac{24}{33,25}\right), \left(\frac{5}{33,25}\right), \\ &\left(\frac{3}{33,25}\right), \left(\frac{3}{33,25}\right), \left(\frac{1}{33,25}\right) = 0,06, 0,18, 0,06, 0,63, 0,03, 0,072, 0,15, 0,09, 0,09, 0,03 \\ &skor \\ &= \left(\frac{0,336}{1,65}\right), \left(\frac{0,566}{1,65}\right), \left(\frac{0,25}{1,65}\right), \left(\frac{0,358}{1,65}\right), \left(\frac{1,302}{1,65}\right), \left(\frac{0,638}{1,65}\right), \left(\frac{0,075}{1,65}\right), \left(\frac{0}{1,65}\right), \left(\frac{0,05}{1,65}\right), \left(\frac{0,02}{1,65}\right) \end{aligned}$$

$$= 0,20, 0,34, 0,15, 0,21, 0,78, 0,38, 0,04, 0, 0,03, 0,01$$

Tabel 4. 18 Matriks normalisasi *Google Maps*

Alternatif	Negative	Skor
Topic 0	0,06	0,20
Topic 1	0,18	0,34
Topic 2	0,06	0,15
Topic 3	0,63	0,22
Topic 4	0,03	0,79
Topic 5	0,72	0,39
Topic 6	0,15	0,05
Topic 7	0,09	0
Topic 8	0,09	0,03
Topic 9	0,03	0,01

Setelah dilakukannya normalisasi pada *Google Maps* perlu dilakukannya pembobotan. Dalam penelitian ini, setiap kriteria diberikan bobot yang sama karena tidak terdapat prioritas khusus antar kriteria. Nilai normalisasi kemudian dikali dengan bobot yaitu 0,5 dikarenakan penelitian ini menggunakan 2 kriteria sehingga untuk memperoleh matriks ternormalisasi terbobot seperti pada rumus berikut sesuai dengan persamaan 2.2:

$$v_{ij} = w_{ij} \times r_{ij} \quad (2.2)$$

Adapun perhitungan matriks normalisasi terbobot adalah sebagai berikut:

Tabel 4. 19 Matriks ternormalisasi terbobot *Google Maps*

Alternatif	Negative	Skor
Topic 0	$0,06 \times 0,5 = 0,03$	$0,20 \times 0,5 = 0,10$
Topic 1	$0,18 \times 0,5 = 0,09$	$0,34 \times 0,5 = 0,17$
Topic 2	$0,06 \times 0,5 = 0,03$	$0,15 \times 0,5 = 0,07$
Topic 3	$0,63 \times 0,5 = 0,31$	$0,22 \times 0,5 = 0,10$
Topic 4	$0,03 \times 0,5 = 0,01$	$0,79 \times 0,5 = 0,39$
Topic 5	$0,72 \times 0,5 = 0,36$	$0,39 \times 0,5 = 0,19$
Topic 6	$0,15 \times 0,5 = 0,07$	$0,05 \times 0,5 = 0,02$
Topic 7	$0,09 \times 0,5 = 0,04$	$0 \times 0,5 = 0$
Topic 8	$0,09 \times 0,5 = 0,04$	$0,03 \times 0,5 = 0,01$
Topic 9	$0,03 \times 0,5 = 0,01$	$0,01 \times 0,5 = 0,01$

Tahapan selanjutnya yaitu melakukan normalisasi pada *Instagram* sebagai berikut:

$$negatif = \sqrt{7^2 + 3^2 + 0^2 + 1^2 + 1^2 + 1^2 + 0^2 + 5^2 + 0^2 + 0^2 + 2^2} = 9,48683$$

$$skor = \sqrt{\begin{matrix} 0,03583^2 + 0,07059^2 + 0,03846^2 + (-0,1)^2 + (-0,0357)^2 \\ + (-0,012)^2 + 0,05495^2 + (-0,0649)^2 \\ + 0,16667^2 + 0^2 + 0,05797 \end{matrix}} = 0,23988$$

$$\begin{aligned} negatif &= \left(\frac{7}{9,48683}\right), \left(\frac{3}{9,48683}\right), \left(\frac{0}{9,48683}\right), \left(\frac{1}{9,48683}\right), \\ &\left(\frac{1}{9,48683}\right), \left(\frac{1}{9,48683}\right), \left(\frac{0}{9,48683}\right), \left(\frac{5}{9,48683}\right), \left(\frac{0}{9,48683}\right), \left(\frac{0}{9,48683}\right), \left(\frac{2}{9,48683}\right) \\ &= 0,74, 0,32, 0, 0,11, 0,11, 0,11, 0, 0,53, 0, 0, 0,21 \\ skor &= \left(\frac{0,03583}{0,23988}\right), \left(\frac{0,07059}{0,23988}\right), \left(\frac{0,03846}{0,23988}\right), \left(\frac{-0,1}{0,23988}\right), \left(\frac{-0,0357}{0,23988}\right), \\ &\left(\frac{-0,012}{0,23988}\right), \left(\frac{0,05495}{0,23988}\right), \left(\frac{-0,0649}{0,23988}\right), \left(\frac{0,16667}{0,23988}\right), \left(\frac{0}{0,23988}\right), \left(\frac{0,05797}{0,23988}\right) \\ &= 0,15, 0,29, 0,16, -0,42, -0,15, -0,05, 0,22, -0,27, 0,69, 0,0, 0,24 \end{aligned}$$

Tabel 4. 20 Matriks normalisasi *Instagram*

Alternatif	Negative	Skor
Topic 0	0,74	0,15
Topic 1	0,32	0,29
Topic 2	0,00	0,16
Topic 3	0,11	-0,42
Topic 4	0,11	-0,15
Topic 5	0,11	-0,05
Topic 6	0,00	0,23
Topic 7	0,53	-0,27
Topic 8	0,00	0,69
Topic 9	0,00	0,00
Topic 10	0,21	0,24

Setelah dilakukannya normalisasi pada *Instagram* perlu dilakukannya pembobotan. Dalam penelitian ini, setiap kriteria diberikan bobot yang sama karena tidak terdapat prioritas khusus antar kriteria. Nilai normalisasi kemudian dikali dengan bobot untuk memperoleh matriks ternormalisasi terbobot:

Tabel 4. 21 Matriks ternormalisasi terbobot *Instagram*

Alternatif	Negative	Skor
Topic 0	$0,74 \times 0,5 = 0,37$	$0,15 \times 0,5 = 0,07$
Topic 1	$0,32 \times 0,5 = 0,16$	$0,29 \times 0,5 = 0,15$
Topic 2	$0 \times 0,5 = 0$	$0,16 \times 0,5 = 0,08$
Topic 3	$0,11 \times 0,5 = 0,05$	$-0,42 \times 0,5 = -0,21$
Topic 4	$0,11 \times 0,5 = 0,05$	$-0,15 \times 0,5 = -0,07$
Topic 5	$0,11 \times 0,5 = 0,05$	$-0,05 \times 0,5 = -0,03$
Topic 6	$0 \times 0,5 = 0$	$0,23 \times 0,5 = 0,11$
Topic 7	$0,53 \times 0,5 = 0,26$	$-0,27 \times 0,5 = -0,14$
Topic 8	$0 \times 0,5 = 0$	$0,69 \times 0,5 = 0,35$
Topic 9	$0 \times 0,5 = 0$	$0 \times 0,5 = 0$
Topic 10	$0,21 \times 0,5 = 0,11$	$0,24 \times 0,5 = 0,12$

4.6.4 Penentuan solusi ideal dan nilai prefensi

Pada titik ini, solusi ideal positif (A^+) yang merupakan topik paling bermasalah dan solusi ideal negatif (A^-) yang merupakan topik paling tidak bermasalah, selain menghitung nilai preferensi untuk masing-masing pilihan. Karena tujuan penelitian ini adalah menentukan prioritas perbaikan, maka kriteria jumlah negatif merupakan jenis *benefit* dan kriteria skor sentimen merupakan jenis *cost*. Berikut merupakan solusi ideal *Google Maps* dari matriks terbobot sesuai dengan persamaan 2.3 dan 2.4:

$$A^+ = (v_1^+, v_2^+, \dots, v_n^+) \quad (2.3)$$

$$A^- = (v_1^-, v_2^-, \dots, v_n^-) \quad (2.4)$$

$$A^+ = \text{Negatif} = \max$$

$$A^+ = \text{Skor} = \min$$

$$\text{negatif} = 0,36$$

$$\text{skor} = 0$$

$$A^+ = (0,36, 0)$$

$$A^- = \text{Negatif} = \min$$

$$A^- = \text{Skor} = \max$$

$$\text{negatif} = 0,01$$

$$\text{skor} = 0,39$$

$$A^- = (0,01, 0,39)$$

Setelah didapatkannya solusi ideal pada *Google Maps*, selanjutnya solusi ideal pada *Instagram* dari matriks terbobot:

$$A^+ = \text{Negatif} = \max$$

$$A^+ = \text{Skor} = \min$$

$$\text{negatif} = 0,37$$

$$\text{skor} = -0,21$$

$$A^+ = (0,37, -0,21)$$

$$A^- = \text{Negatif} = \min$$

$$A^- = \text{Skor} = \max$$

$$\text{negatif} = 0$$

$$\text{skor} = 0,35$$

$$A^- = (0, 0,35)$$

Setelah didapatkan kedua solusi ideal pada *Google Maps* dan *Instagram*, selanjutnya dilakukan perhitungan jarak. Berikut merupakan salah satu contoh perhitungan jarak pada *Google Maps*, Jarak antara alternatif A_i dengan solusi ideal positif dirumuskan sesuai dengan persamaan 2.5 sebagai berikut:

$$D_i^+ = \sqrt{\sum_{j=1}^n (y_i^+ - y_{ij})^2} \quad (2.5)$$

$$\text{Topic 0} = D_i^+ = \sqrt{(0,36 - 0,03)^2 + (0 - 0,102)^2}$$

$$D_i^+ = \sqrt{(0,33)^2 + (-0,102)^2}$$

$$D_i^+ = \sqrt{0,109 + 0,010}$$

$$D_i^+ = \sqrt{0,120}$$

$$D_i^+ = 0,346$$

Jarak antara alternatif A_i dengan solusi ideal negatif dirumuskan sebagai:

$$D_i^- = \sqrt{\sum_{j=1}^n (y_i^- - y_{ij})^2} \quad (2.6)$$

Berikut merupakan salah satu contoh perhitungan jarak pada *Google Maps*:

$$\text{Topic 0} = D_i^- = \sqrt{(0,015 - 0,030)^2 + (0,39 - 0,102)^2}$$

$$D_i^- = \sqrt{(-0,15) + (0,29)^2}$$

$$D_i^- = \sqrt{0,00023 + 0,0853}$$

$$D_i^- = \sqrt{0,0855}$$

$$D_i^- = 0,292$$

Berdasarkan hasil perhitungan jarak *Google Maps* berikut merupakan hasil jarak pada *Google Maps*:

Tabel 4. 22 Hasil jarak *Google Maps*

Alternatif	D+	D-
Topic 0	0,34601	0,29246
Topic 1	0,32019	0,23489
Topic 2	0,33929	0,31843
Topic 3	0,11726	0,41459
Topic 4	0,52397	0
Topic 5	0,1929	0,39985
Topic 6	0,28656	0,37583
Topic 7	0,31573	0,39481
Topic 8	0,31609	0,37974
Topic 9	0,34585	0,38762

Setelah didapatkannya hasil *Google Maps*, selanjutnya dilakukan perhitungan jarak pada *Instagram*:

$$Topic\ 1 = D_i^+ = \sqrt{(0,37 - 0,16)^2 + (-0,21 - 0,15)^2}$$

$$D_i^+ = \sqrt{(0,211)^2 + (-0,356)^2}$$

$$D_i^+ = \sqrt{0,044 + 0,126}$$

$$D_i^+ = \sqrt{0,171}$$

$$D_i^+ = 0,413$$

$$Topic\ 1 = D_i^- = \sqrt{(0 - 0,16)^2 + (0,35 - 0,15)^2}$$

$$D_i^- = \sqrt{(-0,158)^2 + (0,20)^2}$$

$$D_i^- = \sqrt{0,025 + 0,040}$$

$$D_i^- = \sqrt{0,065}$$

$$D_i^- = 0,255$$

Berdasarkan hasil perhitungan jarak *Instagram* berikut merupakan hasil jarak pada *Instagram*:

Tabel 4. 23 Hasil jarak *Instagram*

Alternatif	D+	D-
<i>Topic 0</i>	0,28313	0,45879
<i>Topic 1</i>	0,41337	0,25516
<i>Topic 2</i>	0,46841	0,26723
<i>Topic 3</i>	0,31623	0,55834
<i>Topic 4</i>	0,34345	0,42512
<i>Topic 5</i>	0,36553	0,37623
<i>Topic 6</i>	0,49033	0,23287
<i>Topic 7</i>	0,12827	0,55
<i>Topic 8</i>	0,66714	0
<i>Topic 9</i>	0,42374	0,3474
<i>Topic 10</i>	0,42174	0,24989

Pada tahap ini kedua *platform* sudah mendapatkan masing-masing jarak. Berikut merupakan rumus nilai prefensi sesuai dengan persamaan 2.7:

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad (2.7)$$

Adapun tahapan berikutnya menghitung nilai prefensi *Google Maps* sebagai berikut:

$$Topic\ 0 = \frac{0,292}{0,292 + 0,346} = 0,458$$

$$Topic\ 1 = \frac{0,234}{0,234 + 0,320} = 0,423$$

$$Topic\ 2 = \frac{0,318}{0,318 + 0,339} = 0,484$$

$$Topic\ 3 = \frac{0,414}{0,414 + 0,11} = 0,779$$

$$Topic\ 4 = \frac{0}{0 + 0,523} = 0$$

$$Topic\ 5 = \frac{0,399}{0,399 + 0,192} = 0,779$$

$$Topic\ 6 = \frac{0,375}{0,375 + 0,286} = 0,567$$

$$Topic\ 7 = \frac{0,394}{0,394 + 0,315} = 0,555$$

$$Topic\ 8 = \frac{0,379}{0,379 + 0,316} = 0,545$$

$$Topic\ 9 = \frac{0,387}{0,387 + 0,345} = 0,528$$

Berdasarkan perhitungan diatas, maka didapatkan *ranking* pada masing-masing topik yaitu:

Tabel 4. 24 *Rank* nilai prefensi *Google Maps*

Alternatif	Nilai Prefensi	Rank
<i>Topic 3</i>	0,77952	1
<i>Topic 5</i>	0,67457	2
<i>Topic 6</i>	0,56739	3
<i>Topic 7</i>	0,55565	4
<i>Topic 8</i>	0,54574	5
<i>Topic 9</i>	0,52847	6
<i>Topic 2</i>	0,48414	7
<i>Topic 0</i>	0,45806	8
<i>Topic 1</i>	0,42316	9
<i>Topic 4</i>	0	10

Setelah didapatkannya hasil *Google Maps*, selanjutnya dilakukan perhitungan jarak pada *Instagram*:

$$Topic\ 0 = \frac{0,458}{0,458 + 0,289} = 0,618$$

$$Topic\ 1 = \frac{0,255}{0,255 + 0,413} = 0,381$$

$$Topic\ 2 = \frac{0,267}{0,267 + 0,468} = 0,363$$

$$Topic\ 3 = \frac{0,558}{0,558 + 0,316} = 0,638$$

$$Topic\ 4 = \frac{0,425}{0,425 + 0,343} = 0,553$$

$$Topic\ 5 = \frac{0,376}{0,376 + 0,365} = 0,507$$

$$Topic\ 6 = \frac{0,232}{0,232 + 0,490} = 0,322$$

$$Topic\ 7 = \frac{0,55}{0,55 + 0,128} = 0,810$$

$$Topic\ 8 = \frac{0}{0 + 0,667} = 0$$

$$Topic\ 9 = \frac{0,347}{0,347 + 0,423} = 0,450$$

$$Topic\ 10 = \frac{0,249}{0,249 + 0,421} = 0,372$$

Berdasarkan perhitungan diatas, maka didapatkan *ranking* pada masing-masing topik yaitu:

Tabel 4. 25 *Rank* nilai prefensi *Instagram*

Alternatif	Nilai Prefensi	Rank
<i>Topic 7</i>	0,81089	1
<i>Topic 3</i>	0,63842	2
<i>Topic 0</i>	0,61838	3
<i>Topic 4</i>	0,55314	4
<i>Topic 5</i>	0,50721	5
<i>Topic 9</i>	0,4505	6
<i>Topic 1</i>	0,38167	7
<i>Topic 10</i>	0,37206	8
<i>Topic 2</i>	0,36326	9
<i>Topic 6</i>	0,322	10
<i>Topic 8</i>	0	11

Pada hasil perhitungan *Google Maps* didapatkan *rank* pertama yaitu topik 3 dengan nilai prefensi 0,799 sehingga menjadi prioritas utama yang perlu diperhatikan dan pada topik 4 diurutan terakhir dengan nilai prefensi 0. Pada nilai prefensi *Instagram* didapatkan *rank* pertama yaitu topik 7 dengan nilai prefensi 0,810 sehingga menjadi prioritas utama yang perlu diperhatikan dan topik 8 pada urutan terakhir dengan nilai prefensi 0.

BAB V

PEMBAHASAN

5.1 Perbandingan Hasil Analisis antar *Platform*

Hasil analisis menunjukkan bahwa ulasan pengunjung pada *Instagram* dan *Google Maps* memiliki karakteristik yang berbeda. Perbedaan tersebut terlihat pada hasil *topic modelling*, analisis sentimen, dan prioritas perbaikan menggunakan metode TOPSIS. Perbedaan karakteristik ini dipengaruhi oleh fungsi masing-masing *platform* dalam menyampaikan pengalaman wisata. *Platform* ulasan seperti *Google Maps* lebih banyak dimanfaatkan untuk memberikan evaluasi terhadap pengalaman setelah berkunjung (*post-visit experience*), sedangkan *Instagram* lebih berperan dalam membentuk citra destinasi melalui konten visual dan interaksi media sosial (*pre-visit experience*) sehingga menghasilkan pola opini yang berbeda (Lam dkk., 2020).

Dalam *Google Maps*, ulasan yang diberikan pengunjung cenderung lebih rinci dan deskriptif. Hal ini ditunjukkan oleh banyaknya ulasan yang membahas fasilitas, harga, pelayanan, kenyamanan, serta pengalaman selama berada di lokasi wisata. Selain itu, beberapa ulasan juga memperlihatkan adanya perbandingan antara ekspektasi pengunjung dengan pengalaman yang diterima, terutama berkaitan dengan biaya, fasilitas, dan kenyamanan. Karakteristik tersebut sejalan dengan fungsi *Google Maps* sebagai *platform* ulasan yang mendorong pengguna memberikan evaluasi terhadap pengalaman aktual selama mengunjungi destinasi wisata (Miftahus Sa'adah dkk., 2025).

Sementara itu, ulasan di *Instagram* cenderung lebih singkat, emosional, dan dipengaruhi oleh konten visual yang ditampilkan pada platform tersebut. Topik yang muncul lebih banyak berkaitan dengan *spot* foto, suasana tempat, aktivitas keluarga, hiburan, dan biaya. Hal ini menunjukkan bahwa konten promosi wisata memiliki pengaruh yang lebih besar terhadap persepsi pengguna *Instagram*. Akibatnya, komentar yang muncul sering menunjukkan ketertarikan awal, antusiasme, hingga kekecewaan ketika pengalaman yang diperoleh tidak sesuai dengan citra destinasi yang dibangun melalui media sosial. Temuan ini sejalan dengan penelitian yang menjelaskan bahwa *Instagram* berperan besar dalam membentuk *destination image* dan memengaruhi niat berkunjung melalui kekuatan konten visual dan *user-generated content* (Kusumasondjaja, 2021).

Hasil analisis sentimen juga menunjukkan perbedaan pola pada kedua *platform*. *Google Maps* didominasi ulasan positif pada beberapa aspek seperti kenyamanan, pelayanan, dan makanan sehingga menunjukkan bahwa sebagian besar pengunjung merasa puas terhadap layanan yang diberikan. Namun, masih terdapat ulasan negatif pada aspek tertentu seperti biaya dan pengalaman wisata yang menunjukkan adanya kesenjangan antara ekspektasi dan pengalaman aktual pengunjung. Pada *Instagram*, distribusi sentimen terlihat lebih beragam dan beberapa topik memiliki nilai sentimen negatif, terutama pada aspek harga dan aktivitas keluarga. Kondisi tersebut menunjukkan bahwa pengguna Instagram lebih sensitif terhadap kesesuaian antara konten promosi dengan pengalaman yang diterima. Perbedaan pola sentimen antar *platform* juga ditemukan pada penelitian sebelumnya yang menunjukkan bahwa setiap platform memiliki karakteristik komunikasi dan perilaku pengguna yang berbeda sehingga menghasilkan distribusi sentimen yang tidak sama (Septa dkk., 2026).

Karakteristik ulasan yang berbeda dari masing-masing *platform* memengaruhi hasil prioritas TOPSIS. *Google Maps* lebih menekankan evaluasi berdasarkan pengalaman langsung pengunjung, sedangkan *Instagram* lebih mencerminkan persepsi visual, ekspektasi, dan daya tarik konten digital. Oleh karena itu, penggunaan dua platform dalam penelitian ini mampu memberikan gambaran yang lebih komprehensif mengenai persepsi pengunjung terhadap HeHa Forest. Hasil tersebut kemudian menjadi dasar dalam penyusunan rekomendasi perbaikan layanan menggunakan pendekatan *Lotus Blossom* pada pembahasan selanjutnya (Septa Sintiya dkk., 2026).

5.2 Hasil Prioritas TOPSIS

Hasil perhitungan menggunakan metode TOPSIS menunjukkan bahwa setiap *platform* memiliki prioritas perbaikan yang berbeda berdasarkan persepsi pengunjung terhadap layanan HeHa Forest. Perbedaan tersebut menunjukkan bahwa karakteristik platform digital memengaruhi cara pengunjung mengevaluasi pengalaman wisata. *Google Maps* lebih banyak mencerminkan pengalaman aktual setelah berkunjung, sedangkan *Instagram* lebih dipengaruhi oleh persepsi awal yang terbentuk dari konten visual dan promosi digital. Temuan ini sejalan dengan penelitian yang menyatakan bahwa media sosial dan *platform* ulasan memiliki karakteristik informasi yang berbeda sehingga menghasilkan pola evaluasi wisatawan yang berbeda pula (Guo dkk., 2021).

Salah satu prioritas perbaikan pada *Google Maps* adalah Spot Foto Menarik. Hasil ulasan menunjukkan bahwa meskipun HeHa Forest dikenal memiliki konsep visual dan area foto yang menarik, beberapa pengunjung merasa pengalaman visual yang diperoleh belum sepenuhnya sesuai dengan ekspektasi. Beberapa ulasan menunjukkan bahwa pengunjung mengharapkan *spot* foto yang lebih variatif, kondisi area yang lebih terawat, serta pengalaman visual yang lebih konsisten. Hal tersebut menunjukkan bahwa kualitas elemen visual merupakan salah satu faktor yang memengaruhi kepuasan pengunjung terhadap destinasi wisata. Penelitian sebelumnya juga menjelaskan bahwa daya tarik visual dan estetika destinasi merupakan faktor penting yang memengaruhi pengalaman wisatawan, khususnya pada destinasi yang mengandalkan aktivitas fotografi dan media sosial (Song dkk., 2020).

Prioritas selanjutnya pada *Google Maps* berada pada aspek Makanan dan Taman Bermain. Kondisi tersebut menunjukkan bahwa pengunjung tidak hanya memperhatikan daya tarik visual destinasi, tetapi juga mempertimbangkan pengalaman wisata secara menyeluruh. Aspek makanan mencakup kualitas, cita rasa, variasi menu, serta kesesuaian harga dengan pengalaman yang diterima. Sementara itu, aspek taman bermain menunjukkan bahwa wisatawan keluarga merupakan salah satu segmen penting bagi HeHa Forest sehingga kualitas fasilitas pendukung turut memengaruhi kepuasan pengunjung. Temuan tersebut sejalan dengan penelitian yang menyatakan bahwa kualitas fasilitas, kuliner, dan aktivitas pendukung merupakan faktor yang berkontribusi terhadap kepuasan dan niat berkunjung kembali pada destinasi wisata (Suhartanto, 2021).

Berbeda dengan *Google Maps*, hasil TOPSIS pada Instagram menunjukkan bahwa Biaya Aktivitas Keluarga menjadi prioritas utama. Hasil tersebut menunjukkan bahwa pengguna *Instagram* memiliki sensitivitas yang lebih tinggi terhadap biaya yang harus dikeluarkan selama berwisata. Kondisi ini dapat dipengaruhi oleh ekspektasi yang terbentuk melalui konten visual promosi di media sosial. Konten yang menarik dapat membentuk persepsi bahwa destinasi menawarkan pengalaman wisata yang lebih tinggi. Ketika pengalaman aktual tidak sebanding dengan biaya yang dikeluarkan, pengunjung cenderung memberikan penilaian negatif. Hasil ini sesuai dengan penelitian yang menjelaskan bahwa kesesuaian antara ekspektasi yang dibentuk media sosial dan pengalaman aktual merupakan salah satu faktor utama yang memengaruhi kepuasan wisatawan (Molinillo, 2021).

Selain aspek biaya, topik Minuman dan Wisata dengan Suasana Alam juga menjadi perhatian pada *platform Instagram*. Selama berada di lokasi, pengunjung tidak hanya memperhatikan tampilan visual tempat wisata, tetapi juga pengalaman konsumsi serta suasana lingkungan yang dirasakan. Hal tersebut menunjukkan bahwa pengalaman wisata modern dibentuk oleh kombinasi berbagai elemen, termasuk kualitas makanan dan minuman, suasana lingkungan, serta kenyamanan yang dirasakan selama berkunjung. Penelitian sebelumnya menjelaskan bahwa pengalaman emosional yang diperoleh wisatawan dari suasana destinasi memiliki pengaruh terhadap kepuasan dan persepsi kualitas destinasi (Rather, 2020)

Secara keseluruhan, hasil TOPSIS menunjukkan adanya perbedaan fokus persepsi antar *platform*. *Google Maps* lebih menggambarkan evaluasi berdasarkan pengalaman aktual dan detail selama berkunjung, sedangkan *Instagram* lebih dipengaruhi oleh persepsi visual, ekspektasi awal, dan pengalaman emosional pengunjung. Oleh karena itu, pengelola HeHa Forest perlu mempertimbangkan karakteristik kedua *platform* dalam menyusun strategi pengembangan layanan. Pendekatan tersebut diharapkan mampu menghasilkan perbaikan yang lebih sesuai dengan kebutuhan serta persepsi wisatawan dari berbagai media digital.

5.3 Rekomendasi Perbaikan Menggunakan Pendekatan *Lotus Blossom*

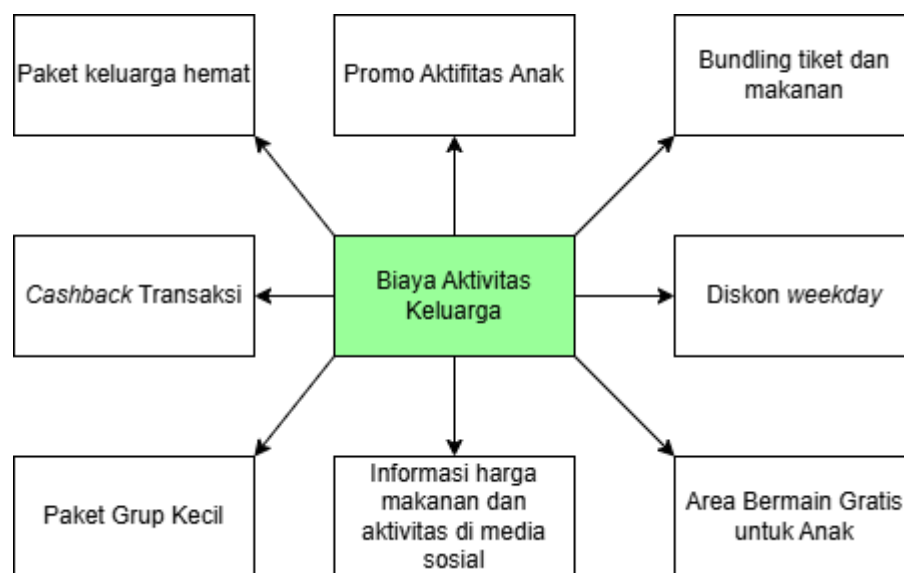
Berdasarkan hasil prioritas yang diperoleh dari metode TOPSIS, pendekatan *Lotus Blossom* digunakan untuk mengembangkan rekomendasi perbaikan secara sistematis berdasarkan topik prioritas dari masing-masing *platform* digital. *Lotus Blossom* merupakan teknik pengembangan ide yang membantu memperluas satu permasalahan utama menjadi beberapa alternatif solusi secara terstruktur sehingga memudahkan proses penyusunan rekomendasi yang lebih komprehensif. Dengan menggunakan pendekatan ini, setiap topik utama dikembangkan menjadi beberapa rekomendasi yang dapat digunakan oleh pengelola HeHa Forest untuk meningkatkan kualitas layanan dan pengalaman pengunjung. Selain itu, pengembangan rekomendasi berdasarkan hasil analisis ulasan pelanggan dinilai mampu menghasilkan strategi perbaikan yang lebih sesuai dengan kebutuhan pengguna karena berasal langsung dari pengalaman dan persepsi pengunjung (Kleinpaul dkk., 2025)

Karakteristik persepsi pengunjung pada masing-masing *platform* menentukan rekomendasi yang dihasilkan. Berdasarkan hasil analisis sebelumnya, prioritas utama

pada platform Instagram adalah Biaya Aktivitas Keluarga, sedangkan pada *Google Maps* adalah Spot Foto Menarik. Perbedaan tersebut menunjukkan bahwa setiap *platform* memiliki fokus evaluasi yang berbeda. Pengguna *Instagram* lebih banyak dipengaruhi oleh ekspektasi yang terbentuk melalui konten visual, sedangkan pengguna *Google Maps* lebih menekankan pengalaman aktual setelah berkunjung (Kleinpaul dkk., 2025). Oleh karena itu, pendekatan *Lotus Blossom* digunakan untuk menghasilkan rekomendasi yang lebih aplikatif, terarah, dan sesuai dengan karakteristik kebutuhan pengunjung pada masing-masing *platform*. Dengan pendekatan tersebut, rekomendasi yang dihasilkan tidak hanya berfokus pada penyelesaian permasalahan utama, tetapi juga mengembangkan alternatif solusi yang saling mendukung sehingga dapat membantu pengelola HeHa Forest dalam meningkatkan kualitas layanan secara berkelanjutan..

5.3.1 Rekomendasi perbaikan Instagram

Hasil TOPSIS pada *platform Instagram* menunjukkan bahwa Biaya Aktivitas Keluarga adalah fokus perbaikan utama. Hasilnya menunjukkan bahwa pelanggan memperhatikan biaya yang mereka habiskan saat melakukan perjalanan bersama keluarga. Oleh karena itu, rekomendasi yang dibuat melalui pendekatan *lotus blossom* difokuskan pada meningkatkan nilai layanan keluarga dan menyesuaikan pengalaman liburan dengan biaya yang dikeluarkan pelanggan.



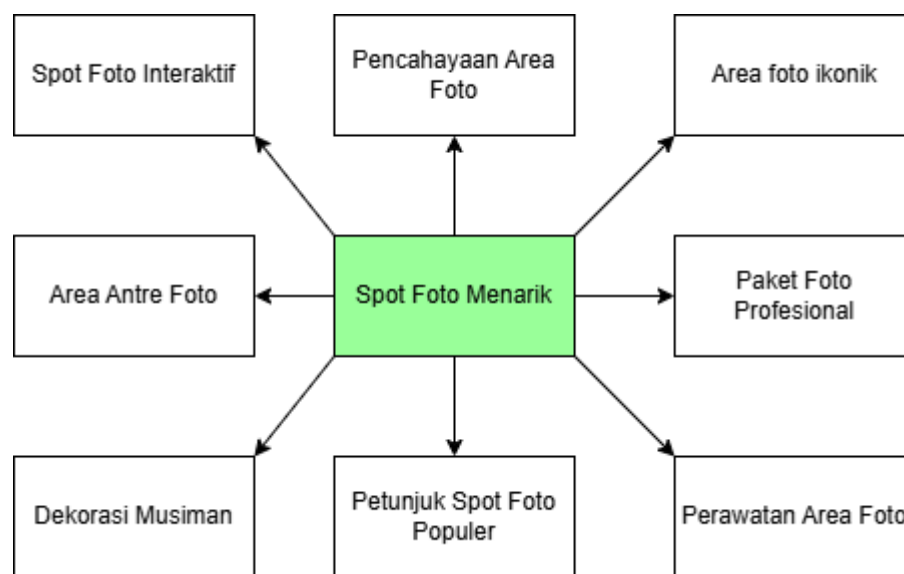
Gambar 5. 1 *Lotus Blossom Instagram*

Berdasarkan gambar 5. 1, beberapa saran yang dibuat termasuk paket keluarga hemat, promo untuk aktivitas anak, paket grup kecil, diskon mingguan, area bermain gratis untuk anak-anak, informasi harga makanan dan aktivitas di media sosial, dan *cashback*

transaksi. Diharapkan rekomendasi ini akan meningkatkan pemahaman pengunjung tentang hubungan antara biaya yang dikeluarkan dan pengalaman liburan yang diterima.

5.3.2 Rekomendasi perbaikan Google Maps

Hasil TOPSIS pada *platform Google Maps* menunjukkan bahwa *Spot Foto Menarik* adalah bagian yang paling penting untuk perbaikan. Hal ini menunjukkan bahwa salah satu faktor utama yang memengaruhi pengalaman pengunjung di *HeHa Forest* adalah kualitas pengalaman visual. Oleh karena itu, rekomendasi yang dibuat melalui pendekatan *lotus blossom* berfokus pada meningkatkan kualitas pengalaman visual dan area *Spot* foto.



Gambar 5. 2 *Lotus Blossom Google Maps*

Berdasarkan gambar 5. 2, beberapa saran yang dibuat termasuk dekorasi musiman, pencahayaan area foto, *Spot* foto interaktif, area antrean, paket foto profesional, rekomendasi *Spot* foto populer, perawatan area foto, dan *Spot* foto ikonik. Diharapkan rekomendasi ini akan meningkatkan daya tarik visual destinasi wisata sekaligus memberikan pengalaman foto yang lebih baik bagi pengunjung.

5.4 Visualisasi Konsep Implementasi Rekomendasi

Berdasarkan hasil rekomendasi yang diperoleh melalui pendekatan *Lotus Blossom*, penelitian ini menyajikan visualisasi konsep implementasi sebagai representasi dari rekomendasi yang dihasilkan. Pada proses *Knowledge Discovery in Databases (KDD)*, visualisasi merupakan tahapan akhir (*knowledge presentation*) yang bertujuan menyampaikan hasil analisis dalam bentuk yang lebih mudah dipahami sehingga dapat

mendukung proses pengambilan keputusan (I. Abrar, 2020). Oleh karena itu, visualisasi pada penelitian ini disusun untuk memberikan gambaran implementasi dari rekomendasi perbaikan yang telah dihasilkan berdasarkan hasil analisis ulasan pengunjung.

Visualisasi yang disajikan bertujuan memberikan gambaran yang lebih jelas mengenai bentuk implementasi rekomendasi sehingga dapat membantu pengelola HeHa Forest dalam memahami dan menerapkan strategi pengembangan yang diusulkan. Visualisasi tersebut bersifat konseptual dan tidak dimaksudkan sebagai rancangan teknis, melainkan sebagai ilustrasi penerapan rekomendasi yang representatif berdasarkan hasil analisis topic modelling, analisis sentimen, metode TOPSIS, dan pendekatan Lotus Blossom. Dengan demikian, hasil penelitian tidak hanya menghasilkan rekomendasi perbaikan, tetapi juga memberikan gambaran implementasi yang dapat menjadi referensi dalam proses pengembangan layanan wisata.

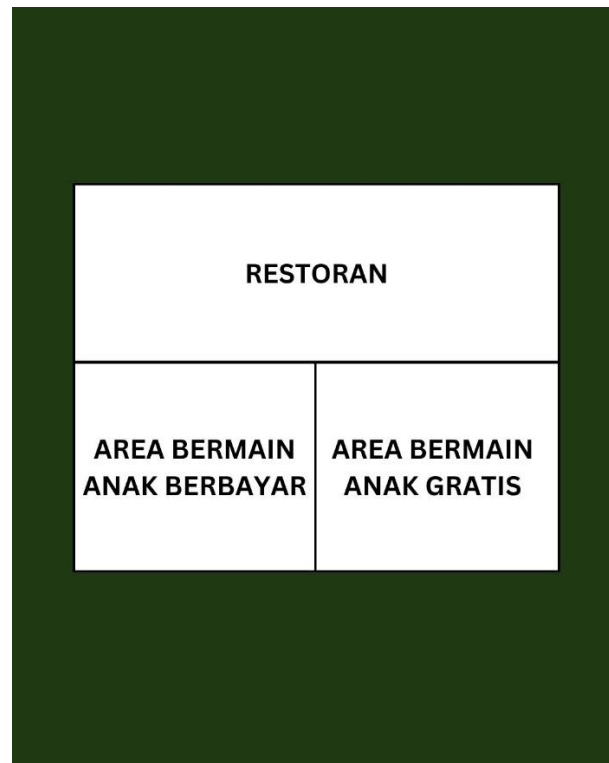
5.4.1 Visualisasi konsep Instagram

Berdasarkan hasil analisis sentimen pada *platform Instagram* terkait atribut informasi harga makanan dan aktivitas di media sosial, diperoleh 77 komentar yang terdiri dari 72 komentar netral dan 5 komentar negatif. Mayoritas komentar netral berupa pertanyaan mengenai harga tiket masuk, biaya wahana, biaya tambahan *Spot* foto, serta fasilitas yang diperoleh pengunjung. Pada komentar negatif menunjukkan adanya persepsi bahwa informasi biaya yang tersedia belum tersampaikan secara lengkap sehingga beberapa pengunjung merasa biaya yang dikeluarkan lebih besar dari ekspektasi. Hasil tersebut menunjukkan bahwa kebutuhan utama pengguna bukan hanya terkait besaran biaya, tetapi juga transparansi informasi harga dan aktivitas yang tersedia sebelum berkunjung. Oleh karena itu, konsep perbaikan difokuskan pada penyediaan informasi harga yang lebih jelas melalui media sosial, penyediaan area bermain gratis untuk anak, serta paket keluarga hemat guna meningkatkan persepsi nilai yang diterima pengunjung. Pada konsep visual informasi harga makanan dan aktivitas pengunjung memiliki 62,3% komentar yang membahas hal tersebut, pada area bermain gratis untuk anak terdapat 7,7% yang membahas hal tersebut, dan pada paket keluarga hemat terdapat 19,4% komentar yang membahas hal tersebut.

INFORMASI HARGA HEHA FOREST	
TIKET MASUK	: Rp. XX.XXX
AREA BERMAIN ANAK	: Rp. XX.XXX
VOUCHER MAKANAN	: Rp. XX.XXX
SPOT FOTO	: Rp. XX.XXX
TIKET PARKIR MOBIL	: Rp. XX.XXX

Gambar 5. 3 Konsep visual Informasi harga makanan dan aktivitas di Media Sosial

Berdasarkan Gambar 5.3, visualisasi informasi harga makanan dan aktivitas di media sosial merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan Biaya Aktivitas Keluarga pada *platform Instagram*. Hasil analisis menunjukkan terdapat 62,3% komentar yang membahas harga tiket masuk, biaya aktivitas, biaya *spot* foto, dan biaya tambahan lainnya. Mayoritas komentar berupa pertanyaan mengenai rincian biaya yang harus dikeluarkan pengunjung sebelum berwisata. Kondisi tersebut menunjukkan bahwa pengunjung membutuhkan informasi biaya yang lebih jelas dan transparan sebelum berkunjung. Oleh karena itu, penyajian informasi harga makanan dan aktivitas melalui media sosial diharapkan dapat membantu pengunjung merencanakan pengeluaran serta mengurangi kesenjangan antara ekspektasi dan pengalaman wisata yang diterima. Penyampaian informasi harga yang transparan dapat meningkatkan kepercayaan pengunjung, mengurangi ketidakpastian sebelum berkunjung, serta membentuk ekspektasi yang lebih realistis. Penelitian oleh Lin dkk. (2023) menunjukkan bahwa transparansi informasi pada platform digital berpengaruh positif terhadap persepsi nilai (*perceived value*) dan kepuasan wisatawan, sehingga penyediaan informasi biaya secara lengkap dapat menjadi strategi untuk meningkatkan kualitas pengalaman pengunjung.



Gambar 5. 4 Konsep visual area bermain gratis untuk anak

Berdasarkan Gambar 5.4, visualisasi area bermain gratis untuk anak merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan Biaya Aktivitas Keluarga pada *platform Instagram*. Hasil analisis ulasan menunjukkan terdapat 7,7% komentar yang secara langsung membahas biaya aktivitas anak, *playground*, dan tiket keluarga. Beberapa pengunjung menyampaikan bahwa aktivitas anak memerlukan biaya tambahan di dalam kawasan wisata sehingga menimbulkan ketidaksesuaian antara ekspektasi dan pengalaman yang diperoleh. Kondisi tersebut menunjukkan bahwa pengunjung keluarga membutuhkan aktivitas yang tetap menarik tanpa menambah biaya tambahan. Oleh karena itu, penyediaan area bermain gratis untuk anak diharapkan dapat meningkatkan nilai manfaat yang diterima keluarga selama berwisata di *HeHa Forest* serta mengurangi kritik terkait aktivitas anak. Penyediaan fasilitas bermain tanpa biaya tambahan dapat meningkatkan persepsi nilai (*value for money*) bagi wisatawan keluarga dan memperpanjang durasi kunjungan. Penelitian Suryaning Endah Yanie dkk. (2024) menjelaskan bahwa fasilitas ramah keluarga merupakan salah satu faktor yang berkontribusi terhadap kepuasan pengunjung dan mendorong niat untuk berkunjung kembali (*revisit intention*).



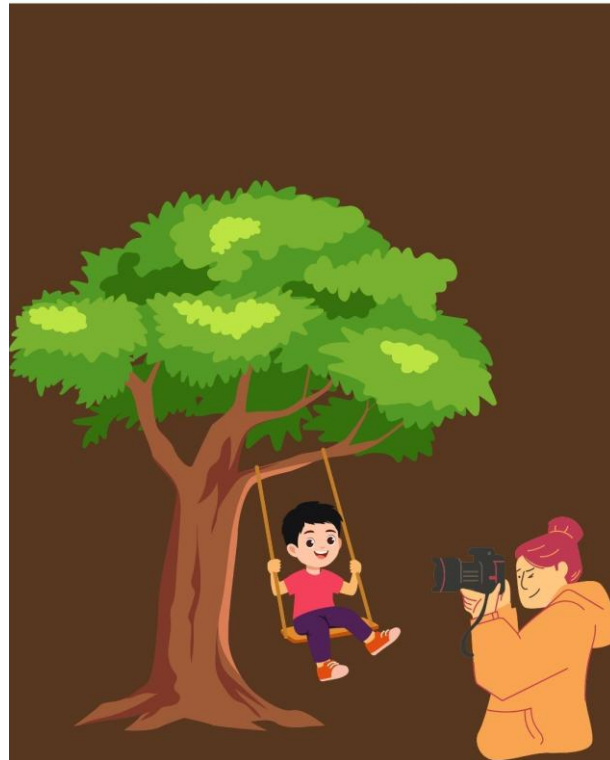
Gambar 5. 5 Paket keluarga hemat

Berdasarkan Gambar 5.5, visualisasi paket keluarga hemat merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan Biaya Aktivitas Keluarga pada *platform Instagram*. Hasil analisis menunjukkan terdapat 19,4% komentar yang mengeluhkan banyaknya biaya tambahan untuk aktivitas, *spot* foto, maupun fasilitas tertentu di dalam kawasan wisata. Beberapa pengunjung juga menyampaikan persepsi bahwa biaya yang dikeluarkan belum sepenuhnya sebanding dengan pengalaman yang diterima. Kondisi tersebut menunjukkan adanya kebutuhan akan pilihan wisata keluarga yang lebih ekonomis dan terjangkau. Oleh karena itu, penyediaan paket keluarga hemat yang menggabungkan tiket masuk dan beberapa aktivitas wisata diharapkan dapat meningkatkan persepsi *value for money* serta memberikan pengalaman wisata keluarga yang lebih baik. Strategi bundling atau paket wisata dapat meningkatkan persepsi keadilan harga (*price fairness*) dan nilai yang diterima wisatawan. Penelitian oleh Setiawan & Susilawati (2025) menunjukkan bahwa persepsi harga yang adil memiliki pengaruh positif terhadap kepuasan pelanggan serta meningkatkan niat untuk berkunjung kembali. Dengan demikian, penerapan paket

keluarga hemat berpotensi meningkatkan daya tarik HeHa Forest bagi segmen wisata keluarga sekaligus memperbaiki persepsi terhadap biaya wisata.

5.4.2 Visualisasi konsep Google Maps

Berdasarkan hasil analisis sentimen pada platform Google Maps terkait atribut *Spot* foto, diperoleh 346 komentar yang terdiri dari komentar positif, netral, dan negatif. Mayoritas komentar positif menunjukkan bahwa pengunjung mengapresiasi banyaknya *Spot* foto yang menarik, suasana yang estetik, serta pemandangan yang mendukung aktivitas fotografi. Namun, masih terdapat komentar netral dan negatif yang menyatakan bahwa beberapa area foto terlihat kurang terawat, variasi *Spot* foto masih terbatas, serta kualitas pencahayaan pada area tertentu belum optimal. Hasil tersebut menunjukkan bahwa *Spot* foto merupakan daya tarik utama HeHa Forest, namun masih terdapat beberapa aspek yang perlu ditingkatkan untuk mendukung pengalaman fotografi pengunjung. Oleh karena itu, konsep perbaikan difokuskan pada pengembangan *Spot* foto interaktif, peningkatan perawatan area foto, serta optimalisasi pencahayaan area foto. Pada konsep visual *Spot* foto interaktif terdapat 63,5% komentar yang membahas *Spot* foto, pada konsep perawatan area foto terdapat 6,9% komentar yang membahas kondisi dan pemeliharaan fasilitas, serta pada konsep pencahayaan area foto terdapat 10,7% komentar yang membahas kualitas pencahayaan dan pengalaman fotografi.



Gambar 5. 6 *Spot* foto interaktif

Berdasarkan Gambar 5.6, visualisasi konsep *Spot* foto interaktif merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan pada atribut aktivitas wisata di platform *Google Maps*. Hasil analisis menunjukkan terdapat 33,5% komentar yang membahas keberadaan, variasi, dan pengalaman penggunaan *Spot* foto di *HeHa Forest*. Mayoritas komentar menyoroti banyaknya *Spot* foto yang tersedia, daya tarik area foto yang *instagramable*, serta harapan pengunjung terhadap *Spot* foto yang lebih beragam dan menarik. Selain itu, beberapa komentar juga menunjukkan bahwa *Spot* foto menjadi salah satu alasan utama pengunjung datang ke *HeHa Forest*. Kondisi tersebut menunjukkan bahwa aktivitas berfoto merupakan daya tarik utama yang memengaruhi pengalaman wisata pengunjung. Temuan ini sejalan dengan penelitian yang menyatakan bahwa pengalaman wisata yang bersifat interaktif dan menyediakan aktivitas yang dapat dibagikan melalui media sosial mampu meningkatkan keterlibatan (*visitor engagement*), kepuasan, serta niat berkunjung kembali karena pengunjung memperoleh pengalaman yang lebih berkesan dan bernilai (Stylidis, 2017). Oleh karena itu, pengembangan *Spot* foto interaktif diharapkan dapat meningkatkan keterlibatan pengunjung, menciptakan pengalaman wisata yang lebih menarik, serta memperkuat daya tarik destinasi sebagai lokasi rekreasi dan fotografi.

Pengecekan Dekorasi	✓
Pembersihan Area	✓
Perbaikan Properti Yang Rusak	✓

Gambar 5. 7 Perawatan area foto

Berdasarkan Gambar 5.7, visualisasi konsep Perawatan area foto merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan pada atribut aktivitas wisata di *platform Google Maps*. Hasil analisis menunjukkan terdapat 6% komentar yang membahas kondisi area wisata dan fasilitas pendukung yang tersedia. Beberapa komentar menyoroti kondisi area foto yang kurang terawat, seperti keberadaan lumut pada kolam dan jalur pejalan kaki, kebersihan beberapa fasilitas, serta kondisi area tertentu yang dinilai belum optimal untuk mendukung kenyamanan pengunjung. Kondisi tersebut menunjukkan bahwa kualitas pengalaman wisata tidak hanya dipengaruhi oleh keberadaan *Spot* foto, tetapi juga oleh kondisi lingkungan dan fasilitas yang mendukung aktivitas pengunjung. Temuan ini sejalan dengan penelitian yang menyatakan bahwa kualitas lingkungan fisik (*physical environment*) dan pemeliharaan fasilitas merupakan faktor penting yang memengaruhi persepsi kualitas destinasi, kepuasan, serta pengalaman wisata secara keseluruhan (Lee, 2011). Oleh karena itu, perawatan area foto dan fasilitas secara berkala diharapkan dapat meningkatkan kenyamanan, keamanan, serta kualitas pengalaman wisata yang dirasakan pengunjung selama berada di HeHa Forest.



Gambar 5. 8 Pencahayaan area foto

Berdasarkan Gambar 5.8, visualisasi konsep pencahayaan area foto merupakan salah satu rekomendasi yang dihasilkan dari pendekatan *Lotus Blossom* berdasarkan prioritas perbaikan pada atribut Spot foto dan pengalaman visual di *platform Google Maps*. Hasil analisis menunjukkan terdapat sekitar 7,7% komentar yang membahas kualitas pencahayaan area foto, khususnya terkait pengalaman berfoto pada malam hari. Sebagian besar komentar menyatakan bahwa tampilan area wisata menjadi lebih menarik ketika lampu-lampu dekoratif dinyalakan sehingga menghasilkan suasana yang lebih estetik dan mendukung aktivitas fotografi. Namun, beberapa komentar juga mengindikasikan bahwa hasil foto pada kondisi tertentu masih kurang optimal, terutama bagi pengunjung yang menggunakan kamera ponsel atau berkunjung pada waktu yang kurang tepat. Kondisi tersebut menunjukkan bahwa pencahayaan memiliki peran penting dalam membentuk kualitas pengalaman visual pengunjung. Penelitian terdahulu juga menjelaskan bahwa elemen visual seperti pencahayaan, suasana, dan estetika lingkungan berkontribusi terhadap pengalaman wisata yang lebih positif, meningkatkan kepuasan pengunjung, serta memperkuat persepsi daya tarik destinasi. Oleh karena itu, peningkatan pencahayaan pada area foto diharapkan dapat memperkuat daya tarik visual destinasi,

menghasilkan kualitas foto yang lebih baik, serta meningkatkan kepuasan pengunjung saat melakukan aktivitas fotografi.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka diperoleh beberapa Kesimpulan sebagai berikut:

1. Dengan menggunakan metode *Latent Dirichlet Allocation* (LDA), penelitian ini berhasil mengidentifikasi topik utama dalam ulasan pengunjung HeHa Forest. Hasil pemodelan topik menunjukkan bahwa karakteristik ulasan pada *Instagram* dan *Google Maps* memiliki perbedaan. Pada *platform Instagram*, topik yang muncul lebih banyak berkaitan dengan *Spot* foto, suasana wisata, aktivitas keluarga, hiburan, dan aspek visual destinasi wisata. Sementara itu, pada *Google Maps*, topik yang muncul lebih banyak berkaitan dengan fasilitas, makanan, pelayanan, kenyamanan, serta pengalaman langsung pengunjung selama berada di lokasi wisata
2. Hasil analisis sentimen yang dilakukan menggunakan metode *InSet Lexicon* menunjukkan bahwa ulasan pengunjung didominasi oleh sentimen positif pada kedua *platform*. Meskipun demikian, ada beberapa topik dengan sentimen negatif, terutama terkait dengan biaya aktivitas keluarga, pengalaman wisata, dan beberapa fasilitas pendukung. Hasil tersebut menunjukkan bahwa masih ada perbedaan antara apa yang diharapkan pengunjung dari HeHa Forest dan apa yang mereka dapatkan.
3. Hasil perhitungan metode TOPSIS menunjukkan bahwa, berdasarkan persepsi pengunjung, setiap *platform* memiliki prioritas perbaikan yang berbeda. Pada *Google Maps*, prioritas perbaikan berfokus pada peningkatan aspek *Spot* Foto Menarik karena visual sangat penting untuk pengalaman wisata pengunjung. Sementara pada *platform Instagram* berfokus pada Biaya Aktivitas Keluarga karena pengunjung sangat peduli dengan bagaimana menyesuaikan antara biaya yang dikeluarkan dengan pengalaman wisata yang diterima.
4. Berdasarkan hasil prioritas permasalahan, penelitian ini berhasil merumuskan rekomendasi strategi pengembangan wisata menggunakan pendekatan *lotus blossom*. Pendekatan tersebut membantu mengembangkan hasil analisis menjadi rekomendasi yang lebih aplikatif dan terstruktur sesuai dengan karakteristik masing-masing *platform*. Rekomendasi yang dihasilkan difokuskan pada peningkatan pengalaman pengunjung, kualitas layanan, serta pengembangan daya tarik wisata HeHa Forest berdasarkan persepsi pengunjung pada *Google Maps* dan *Instagram*.

6.2 Saran

Berdasarkan hasil analisis dan kesimpulan, maka saran yang diberikan dalam penelitian ini adalah sebagai berikut:

1. Bagi HeHa *Forest*

Pengelola HeHa *Forest* disarankan untuk meningkatkan kualitas layanan dengan fokus pada menyesuaikan antara biaya yang dikeluarkan dengan pengalaman wisata yang diterima. Pengelola juga dapat meningkatkan kenyamanan area wisata, menambah aktivitas yang lebih interaktif, dan meningkatkan kualitas *Spot* foto sebagai salah satu daya tarik utama destinasi wisata. Selain itu, elemen pelayanan yang telah mendapatkan persepsi positif dari pengunjung harus dipertahankan.

2. Bagi Universitas

Universitas diharapkan dapat mendorong pengembangan penelitian berbasis data *digital* seperti analisis ulasan *online* untuk memahami perilaku konsumen dan persepsi masyarakat. Selain itu, pembelajaran tentang *topic modelling*, analisis sentimen, dan teknik pengambilan keputusan multikriteria seperti TOPSIS harus terus dikembangkan agar mahasiswa memiliki kemampuan dalam analisis data yang lebih baik dan relevan dengan kemajuan teknologi saat ini.

3. Bagi peneliti selanjutnya

Penelitian selanjutnya disarankan Untuk mendapatkan pemahaman yang lebih baik tentang konteks bahasa yang lebih kompleks, penelitian lebih lanjut harus menggunakan analisis sentimen berbasis *machine learning* atau pendekatan *hybrid*. Untuk meningkatkan interpretasi topik, juga dapat dipikirkan untuk menggunakan metode model topik tambahan seperti *BERTopic*. Untuk membuat hasil penelitian lebih realistis, penelitian juga dapat dikembangkan dengan menambahkan sumber data dari platform digital lain. Selain itu, kriteria yang lebih objektif dapat dibuat dengan menggunakan metode pembobotan seperti *entropy* atau *Analytical Hierarchy Process* (AHP).

DAFTAR PUSTAKA

- Akten, S. (2025). Evaluation Of Demre District Sustainable Tourism Strategies Using SWOT And TOPSIS Analysis. *Sustainability (Switzerland)*, 17(15). <https://doi.org/10.3390/Su17156831>
- Badan Pusat Statistik. (2026). *Badan Pusat Statistik Indonesia*. (3 Maret 2026). *Jumlah Perjalanan Wisatawan Nasional, 2026*. Diakses Pada 13 Maret 2026, Dari <https://www.bps.go.id/id/statistics-table/2/mje5nsmy/jumlah-perjalanan-wisatawan-nasional.html>.
- Badan Pusat Statistik Provinsi DI Yogyakarta. (2026). *Jumlah Kunjungan Wisatawan Nusantara Menurut Tujuan*, 2026. Diakses Pada 16 Maret 2026, Dari <https://yogyakarta.bps.go.id/id/statistics-table/2/nTE0IzI=/jumlah-kunjungan-wisatawan-nusantara-menurut-tujuan-.html>.
- Bigne, E., Ruiz, C., Perez-Cabañero, C., & Cuenca, A. (2023). Are Customer Star Ratings And Sentiments Aligned? A Deep Learning Study Of The Customer Service Experience In Tourism Destinations. *Service Business*, 17(1), 281–314. <https://doi.org/10.1007/S11628-023-00524-0>
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A Comprehensive Survey On Sentiment Analysis: Approaches, Challenges And Trends. *Knowledge-Based Systems*, 226, 107134. <https://doi.org/10.1016/J.KNOSYS.2021.107134>
- Blanco-Moreno, S., González-Fernández, A. M., Muñoz-Gallego, P. A., & Casaló, L. V. (2024a). Understanding Engagement With Instagram Posts About Tourism Destinations. *Journal Of Destination Marketing And Management*, 34. <https://doi.org/10.1016/J.Jdmm.2024.100948>
- Blanco-Moreno, S., González-Fernández, A. M., Muñoz-Gallego, P. A., & Casaló, L. V. (2024b). Understanding Engagement With Instagram Posts About Tourism Destinations. *Journal Of Destination Marketing And Management*, 34. <https://doi.org/10.1016/J.Jdmm.2024.100948>
- Chen, T. C. T., Wu, H. C., & Hsu, K. W. (2024). Recommending Suitable Hotels To Travelers In The Post-COVID-19 Pandemic Using A Novel FAHP-Fuzzy TOPSIS Approach. *Complex And Intelligent Systems*, 10(5), 6901–6915. <https://doi.org/10.1007/S40747-024-01521-0>
- Dwi Ayunda, L., & Andiyansari, P. (2024). *Strategi Komunikasi Pemasaran Digital Heha Forest Dalam Meningkatkan Kunjungan Wisatawan* (Vol. 4, Nomor 1).
- Ekeocha, D. O., Ogbuabor, J. E., Orji, A., & Kalu, U. I. (2021). International Tourism And Economic Growth In Africa: A Post-Global Financial Crisis Analysis. *Tourism Management Perspectives*, 40, 100896. <https://doi.org/10.1016/J.TMP.2021.100896>
- Guo, X., Pesonen, J., & Komppula, R. (2021). Comparing Online Travel Review Platforms As Destination Image Information Agents. *Information Technology And Tourism*, 23(2), 159–187. <https://doi.org/10.1007/S40558-021-00201-W>

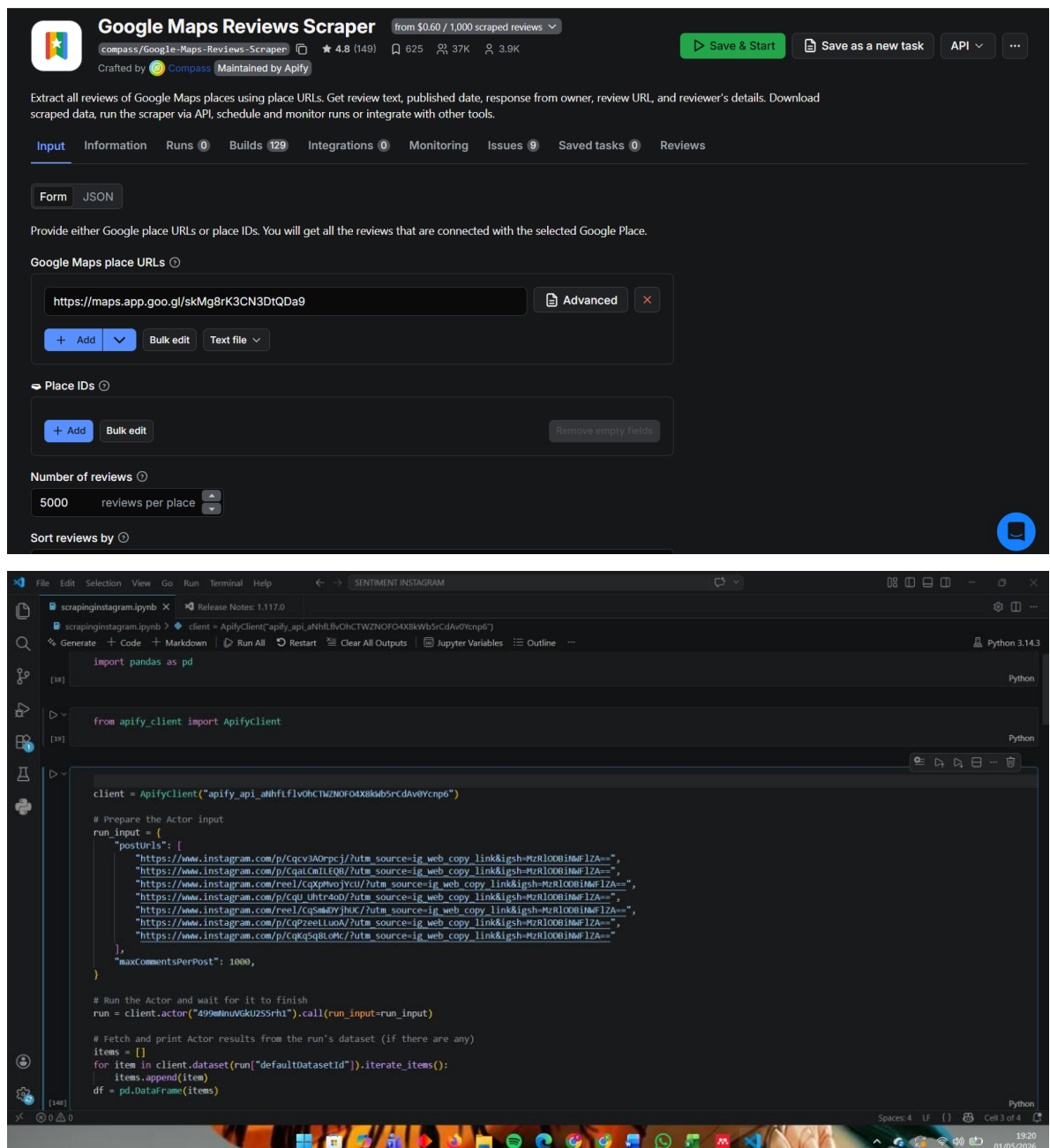
- Hasibuan, I. M., Mutthaqin, S., Erianto, R., & Harahap, I. (T.T.). *KONTRIBUSI SEKTOR PARIWISATA TERHADAP PEREKONOMIAN NASIONAL*. <https://doi.org/10.30651/Jms.V8i2.19280>
- Husnul Khotimah, T., Nilam Novita Sari, Khaola Rachma Adzima, & Leny Dhianti Haeruman. (2025). Pemodelan Topik Pada Analisis Sentimen Terhadap Pendidikan Literasi Numerasi Di Indonesia Menggunakan Latent Dirichlet Allocation. *JURNAL RISET PEMBELAJARAN MATEMATIKA SEKOLAH*, 9(2), 9–20. <https://doi.org/10.21009/Jrpms.092.02>
- I. Abrar, Z. A. F. M. And A. M. B. (2020). A Machine Learning Approach For Intrusion Detection System On NSL-KDD Dataset. *2020 International Conference On Smart Electronics And Communication (ICOSEC), Trichy, India, 2020, Pp. 919-924*.
- Jeong, H. (2025). Exploring Research Trends In Thai Learners Studying Korean Via Topic Modeling And Keyword Network Analysis. *Journal Of Arts And Thai Studies*, 47(1). <https://doi.org/10.69598/Artssu.2025.4039>
- Jiang, T., Xu, Y., Li, Y., & Xia, Y. (2025). Integration Of Public Libraries And Cultural Tourism In China: An Analysis Of Library Attractiveness Components Based On Tourist Review Mining. *Information Processing & Management*, 62(2), 104000. <https://doi.org/10.1016/J.IPM.2024.104000>
- Lee, S. , J. S. , & K. D. (2011). *The Impact Of Tour Quality And Tourist Satisfaction On Tourist Loyalty*.
- Kirilenko, A. P., & Stepchenkova, S. (2025). Facilitating Topic Modeling In Tourism Research: Comprehensive Comparison Of New AI Technologies. *Tourism Management*, 106, 105007. <https://doi.org/10.1016/J.TOURMAN.2024.105007>
- Kleinpaul, E., Pires, A., Machado, L., & Parizi, R. (2025). *On The Use Of The Lotus Blossom Technique To Foster Collaboration Among Students In Software Development*.
- Koto, F., & Rahmaningtyas, G. Y. (2017). Inset Lexicon: Evaluation Of A Word List For Indonesian Sentiment Analysis In Microblogs. *Proceedings Of The 2017 International Conference On Asian Language Processing, IALP 2017, 2018-January*, 391–394. <https://doi.org/10.1109/IALP.2017.8300625>
- Kristina, T. (2018). *Sistem Pendukung Keputusan Dengan Menggunakan Metode TOPSIS*. XX(1).
- Kurniawan, K., & Louvan, S. (2018). Indosum: A New Benchmark Dataset For Indonesian Text Summarization. *Proceedings Of The 2018 International Conference On Asian Language Processing, IALP 2018*, 215–220. <https://doi.org/10.1109/IALP.2018.8629109>
- Kusumasondjaja, S. (2021). Building Social Media Engagement On Instagram By Using Visual Aesthetics And Message Orientation Strategy: A Content Analysis On Instagram Content Of Indonesia Tourism Destinations. *Journal Of International Conference Proceedings*, 4(3). <https://doi.org/10.32535/Jicp.V4i3.1304>

- Lam, J. M. S., Ismail, H., & Lee, S. (2020). From Desktop To Destination: User-Generated Content Platforms, Co-Created Online Experiences, Destination Image And Satisfaction. *Journal Of Destination Marketing & Management*, 18, 100490. <https://doi.org/10.1016/J.JDMM.2020.100490>
- Lin, H. C., Liu, X., Huang, Y., & Chen, H. Y. (2023). Determinants Of Continued Use Of Tourism And Hospitality E-Commerce Platforms And The Role Of Information Transparency. *Current Issues In Tourism*, 26(19), 3140–3159. <https://doi.org/10.1080/13683500.2022.2147052>
- Mao, Z. X., Zhu, Y. R., Hou, J. P., Ma, C. P., & Liu, S. P. (2021). Multiple Diamond-Alpha Integral In General Form And Their Properties, Applications. *Mathematics*, 9(10). <https://doi.org/10.3390/Math9101123>
- Mayasari, R., Fithiasari, K., & Prastyo, D. (2021, Oktober 19). *Text Mining To Analyse Publication Topics Of COVID-19 Using HDP And LDA Methods*. <https://doi.org/10.4108/Eai.19-12-2020.2309174>
- Mensah, E. A., Odame, D. N. A., Ankrah, I., Obuobisa-Darko, T., & Hinson, R. E. (2025). From Reviews To Reflections: Understanding Tourist Sentiments And Satisfaction In African Destinations Through User-Generated Content. *Annals Of Tourism Research Empirical Insights*, 6(1), 100174. <https://doi.org/10.1016/J.ANNALE.2025.100174>
- Merawati, N. L. P., Amrullah, A. Z., & Ismarmiaty. (2021). Analisis Sentimen Dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes Dan Latent Dirichlet Allocation. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 123–131. <https://doi.org/10.29207/RESTI.V5I1.2587>
- Miftahus Sa'adah, S., Umam, K., Handayani, M. R., & Mustofa, I. (2025). Mapping The Polarity Of Tourist Opinions On Indonesian Destinations Through Google Maps Reviews Using Supervised Learning Methods. *Journal Of Applied Informatics And Computing (JAIC)*, 9(5), 2845–2853. <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Ramdani, R., & Cahyana, R. (2025). Analisis Sentimen Ulasan Wisata Budaya Menggunakan Metode Support Vector Machine Dan Long Short-Term Memory. *Jurnal Algoritma*, 22(2), 1188-1199–1188–1199. <https://doi.org/10.33364/ALGORITMA/V.22-2.2585>
- Rather, R. & S. J. (2020). *Influence Of Customer Experience On Memories And Loyalty In Tourism And Hospitality Industry., GITAM Journal Of Management*.
- Sa'adah, S. M., Umam, K., Handayani, M. R., & Mustofa, M. I. (2025). Mapping The Polarity Of Tourist Opinions On Indonesian Destinations Through Google Maps Reviews Using Supervised Learning Methods. *Journal Of Applied Informatics And Computing*, 9(5), 2845–2853. <https://doi.org/10.30871/JAIC.V9I5.9836>
- Septa Sintiya, E., Azizah Rendra Putri, S., Budi Prasetyo, G., Bella Vista Teknologi Informasi, C., Negeri Malang Jl Soekarno Hatta No, P., Lowokwaru, K., & Malang, K. (2026). ANALISIS SENTIMEN MULTI-PLATFORM UNTUK STRATEGI

- PEMASARAN DIGITAL PADA DESTINASI WISATA. Dalam *Jurnal Mahasiswa Teknik Informatika* (Vol. 10, Nomor 3).
- Setiawan, R., & Susilawati, W. (2025). Revisit Intention Through The Development Of Tourism Components And Tourist Satisfaction At Hot Spring Tourism Destinations. *Khazanah Sosial*, 7(1), 16–31. <https://doi.org/10.15575/KS.V7I1.38996>
- Singgale, Y. A. (2023). Analisis Sentimen Pengunjung Pulau Komodo Dan Pulau Rinca Di Website Tripadvisor Berbasis CRISP-DM. *Journal Of Information System Research (JOSH)*, 4(2), 614–625. <https://doi.org/10.47065/JOSH.V4I2.2999>
- Song, P., Guo, J., Xu, E., Mayer, A. L., Liu, C., Huang, J., Tian, G., & Kim, G. (2020). Hydrological Effects Of Urban Green Space On Stormwater Runoff Reduction In Luohe, China. *Sustainability (Switzerland)*, 12(16). <https://doi.org/10.3390/Su12166599>
- Stylidis, D. , S. A. , & B. Y. (2017). *Destination Image, On-Site Experience And Behavioural Intentions*.
- Suárez-Rojas, C., Lam-González, Y. E., & Martín, J. C. (2025). Not So Demanding! Employing The Fuzzy-Hybrid TOPSIS To Explain (Un)Demanding Whale-Watcher Behaviour. *Journal Of Outdoor Recreation And Tourism*, 51, 100919. <https://doi.org/10.1016/J.JORT.2025.100919>
- Suhartanto, D. , D. D. , I. T. A. T. , & S. R. (2021). *Mobile Food Ordering Adoption In Tourism And Hospitality*.
- Suryaning Endah Yanie, D., Artanti, Y., Program Studi, M. S., Fakultas Ekonomika Dan Bisnis Dosen Program Studi, M. S., & Fakultas Ekonomika Dan Bisnis Universitas Negeri Surabaya, M. (2024). *PENGARUH PERSEPSI NILAI TERHADAP KEPUASAN WISATAWAN PADA KEBUN RAYA MANGROVE (KRM) SURABAYA THE INFLUENCE OF PERCEIVED VALUE ON TOURIST SATISFACTION AT KEBUN RAYA MANGROVE (KRM) SURABAYA*.
- Tiwari Paras, Tripathi Ashutosh, Singh Avaneesh, & Rai Sawan. (2023). *Advanced Hierarchical Topic Labeling For Short Text*. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10092802>
- Wang, X., Yang, Z., & Guo, Y. (2024). Research On The Influencing Factors Of Cultural And Tourism Service Quality In Huizhou Area. *Sustainability (Switzerland)* , 16(13). <https://doi.org/10.3390/Su16135535>
- Wu, D. C., Zhong, S., Song, H., & Wu, J. (2024). Do Topic And Sentiment Matter? Predictive Power Of Online Reviews For Hotel Demand Forecasting. *International Journal Of Hospitality Management*, 120, 103750. <https://doi.org/10.1016/J.IJHM.2024.103750>
- Wulan Sari, A., Iman Hermanto, T., & Defriani, M. (2023). Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, 7(3), 1366–1378. <https://doi.org/10.33395/SINKRON.V8I3.12447>

LAMPIRAN

Lampiran 1 Pengumpulan data



The image displays two screenshots related to data collection using the Apify Google Maps Reviews Scraper.

Top Screenshot: Google Maps Reviews Scraper Interface

- Header:** "Google Maps Reviews Scraper" with a rating of 4.8 (149) and 625 reviews. It shows a limit of "from \$0.60 / 1,000 scraped reviews".
- Buttons:** "Save & Start", "Save as a new task", "API", and a menu icon.
- Input Section:**
 - Toggle between "Form" and "JSON".
 - Text: "Provide either Google place URLs or place IDs. You will get all the reviews that are connected with the selected Google Place."
 - Google Maps place URLs:** A text input field containing "https://maps.app.goo.gl/skMg8rK3CN3DtQDa9". Below it are buttons for "+ Add", "Bulk edit", and "Text file".
 - Place IDs:** A section with "+ Add", "Bulk edit", and "Remove empty fields" buttons.
 - Number of reviews:** A dropdown menu set to "5000 reviews per place".
 - Sort reviews by:** A dropdown menu.

Bottom Screenshot: Jupyter Notebook

The notebook shows the configuration and execution of the scraper:

```

import pandas as pd

from apify_client import ApifyClient

client = ApifyClient("apify_api_ahft1vOhCTWZNOF04XBWb5rCdAv0Ycnp6")

# Prepare the Actor input
run_input = {
    "postUrls": [
        "https://www.instagram.com/p/Cqcv3A0rpij/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/p/CqalCm1EQ8/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/reel/CqXpVojYcu/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/p/CqUUhtr4oD/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/reel/CqS6dYVjhUC/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/p/CqPzeelLuoA/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA==",
        "https://www.instagram.com/p/CqKq5q8LoMc/?utm_source=ig_web_copy_link&igsh=MzRlODBiNmFlZA=="
    ],
    "maxCommentsPerPost": 1000,
}

# Run the Actor and wait for it to finish
run = client.actor("499mmuVgKu255rh1").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)
  
```

2

Untitled11.ipynb

File Edit View Insert Runtime Tools Help

Commands Code Text Run all

Connect

```

import pandas as pd
import re
import nltk

# Install sastrawi if not already installed
!pip install sastrawi

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory

# Pastikan NLTK data 'punkt' terunduh
try:
    # Attempt to find the specific 'punkt.tab' resource directly
    nltk.data.find('tokenizers/punkt_tab')
except LookupError:
    # If 'punkt.tab' is not found, download it as suggested by the error message
    nltk.download('punkt_tab')

# --- 1. Muat Dataset ---
try:
    df = pd.read_excel('sentiment_maps_update(21).xlsx')
    print('Dataset berhasil dimuat.')
except FileNotFoundError:
    print('Error: File sentiment_maps_update(21).xlsx tidak ditemukan. Pastikan file sudah diunggah.')
    exit()
except Exception as e:
    print(f'Terjadi error saat memuat dataset: {e}')
    exit()

```

Variables Terminal

19:23 01/05/2023

```

# Tampilkan nama kolom asli untuk verifikasi
print("Nama-nama kolom asli yang tersedia:", df.columns.tolist())

# --- 2. Pemilihan Kolom Awal dan Persiapan Data ---
# kolom 'title' hingga 'indonesia_translate' dibiarkan
# kolom 'sentiment score' dan 'sentiment' dibiarkan
# kolom lain akan dibuat ulang atau disesuaikan
base_cols = df.loc[:, 'title':'indonesia_translate'].columns.tolist()
sentiment_cols = ['sentiment score', 'sentiment']
existing_cols_to_keep = base_cols + sentiment_cols

# Buat DataFrame baru hanya dengan kolom yang akan dipertahankan atau sebagai sumber
df_processed = df[existing_cols_to_keep].copy()

# Pastikan kolom 'indonesia_translate' tidak kosong dan bertipe string
df_processed['indonesia_translate'] = df_processed['indonesia_translate'].astype(str).fillna('')

# --- 3. Preprocessing Teks ---

# a. Case Folding
print("Melakukan Case folding...")
df_processed['case_folding'] = df_processed['indonesia_translate'].apply(lambda x: str(x).lower())

# --- Stopword Removal (on string, before tokenizing) ---
print("Melakukan Stopword Removal...")

# Inisialisasi StopwordRemover dari Sastrawi

```

```

# Inisialisasi StopwordRemover dari Sastrawi
factory_stop = StopwordRemoverFactory()
stopwords_sastrawi = factory_stop.get_stop_words()

# Tambahkan custom stopwords
custom_stopwords = ['yang', 'dan', 'dengan', 'ini', 'itu', 'di', 'ke', 'dari', 'untuk',
'pada', 'sangat', 'nya', 'saya', 'banget', 'ke', 'ini', 'ya', 'aku', 'mau', 'di', 'dan', 'kesini', 'nya',
'bgt', 'yg', 'yang', 'ada', 'kak', 'ga', 'bisa', 'nih', 'buat', 'juga',
'bagu', 'jadi', 'tempatny', 'min', 'sih', 'lagi', 'gak', 'kesana', 'tempat',
'aja', 'karena', 'ikutin', 'masuk', 'kalu', 'itu', 'sini', 'sama', 'tapi',
'pernah', 'semoga', 'banyak', 'belum', 'seru',
# Additional common Indonesian stopwords not in the top 50
'untuk', 'dengan', 'dari', 'oleh', 'pada', 'dalam', 'akan', 'adalah',
'tidak', 'sangat', 'dan', 'atau', 'hanya', 'kita', 'dengan', 'akan', 'juga', 'ada', 'mereka', 'sudah', 'saya', 'ter',
'wow', 'hehe', 'ya', 'ayo', 'oh', 'bro', 'dong', 'nih', 'deh', 'banget', 'bgt',
'keren', 'bagus', 'mantap', 'sip', 'wak', 'senang', 'kecewa', 'hahaha',
'tenan', 'baik', 'santa', 'bagus', 'jadi', 'saja', 'semua', 'ada', 'dapat',
'lebih', 'hehehehe', 'hari', 'malam', 'bulan', 'jau', 'saat', '-',
'photo', 'tamer', 'percaya', 'bukti', 'datang', 'baru', 'jadi', 'dasar',
'pasang', 'cukup', 'sedia', 'pilih', 'lihat',
'dekat', 'jauh', 'tengah', 'gas', 'gasss', 'yok', 'yokkk',
'senang', 'baik', 'cokok', 'rekomendasi', 'melu', 'mas', 'sisi',
'25', '20', 'pol', 'gassssssss', '1', 'jadi', 'satu', 'orang', 'kamu', 'coba', 'ada', 'buat', 'dapat', 'pilih', 'datang', 'baru', 'akhir',
'waktu', 'hari', 'siang', 'dekat', 'jauh', 'sembuh', 'tabrak', 'cukup',
'terlalu', 'sekali', 'per', 'guna', 'buat', 'semua', 'tetap', 'paling', 'cari',
'baik', 'senang', 'rekomendasi', 'benar-benar', 'akhir', 'lama',
'biasa', 'besar', 'kecil', 'kurang', 'atas', 'bosan', 'tinggal', 'unjung', 'kruramah', 'sesal',

```

```

'terlalu', 'sekali', 'per', 'guna', 'buat', 'semua', 'tetap', 'paling', 'cari',
'baik', 'senang', 'rekomendasi', 'benar-benar', 'akhir', 'lama',
'biasa', 'besar', 'kecil', 'kurang', 'atas', 'bosan', 'tinggal', 'unjung', 'kruramah', 'sesal',
'sana', 'kelinci', 'sama', 'santa', 'nikmat', 'idul', 'fitri', 'sendiri', 'kena', 'teduh', 'jangkau',
'kurang', 'mes', 'perlu', 'mungkin', 'tunggu', 'sembuh', 'buka', 'kasih',
'luan', 'tambah', 'salah', 'siap', 'layak', 'tun', 'bantu', 'kerja', 'standar',
'utama', 'senang', 'pas', 'cuaca', 'panas', 'deras', 'vip', 'hwa', 'sukses', 'lezat', 'buruk',
'lihat', 'beri', 'besar', 'coba', '25', 'pes', 'imat', 'ganteng', 'lila', 'stroberi', 'jeep', 'ocean', 'sky',
'film', 'to', 'nah', 'buat', 'semangat', 'rezeki', 'berkah', 'takut', 'sukses', 'wah',
'umayan', 'akhir', 'ada', 'jadi', 'diri', 'istimewa', 'tentang', 'cara', 'temu', 'lupa',
'pergi', 'sedia', 'gratis', 'hidup', 'terobos', 'bangun', 'ajar', 'bikin',
'bicara', 'kenal', 'misi', 'pulang', 'simpan', 'kirin', 'pertamata',
'setengahsetengah', 'hasil', 'soal', 'arti', 'promosi', 'komentar', 'kategori', 'surabaya', 'panjang', 'ulang', 'daftar', 'jamin',
'rekomendasi', 'buat', 'sama', 'jangan', 'tambah', 'dalam', 'bagi', '30', '5', '8410', 'nikmat', 'ada', 'beri', 'jadi', 'lihat', 'suejkkkk', 'pool', 'hargobangun', 'tekan', '8410', 'padu', 'indo',
'berrrrrrr', 'ahhhhh', 'transaksi', 'menyenangkan', '3', 'mes', 'akhir',
'bagi', 'guna', 'digunakan', 'menggunakan', 'instagramable', 'suka', 'puas',
'pikir', 'baik', 'lihat', 'seperti', 'utama', 'kalau', 'bakkan', 'meski',
'benar', 'sesuai', 'harap', 'habis', 'kasi', 'dapat', 'bagian', '0', 'bersama',
'didapatkan', 'dapatkan', 'sedikit', 'meter', 'kawasan', 'jumpa', 'konsep', 'lengkap', 'aman', 'hibur',
'sulit', 'mulai', 'didapat', 'dibawah', 'kebahah', 'disamakan', 'dapat', 'menit', 'sayang',
'lama', 'cari', 'jelas', 'lewat', 'objek', 'baah', 'kumpul', 'rumah', '200',
'bagi', 'sama', 'semuanya', 'terakhir', 'sungguh', 'memang', 'kasih',
'bersedia', 'pohon', 'masuk', 'lihat', 'baru', 'sembah',
'ooo', 'to', 'berbagi', 'memesan', 'terbaru', 'berguna', 'berakhir', 'pendapat',
'mendapatkan', 'terdapat', 'didapatkan', 'bagikan', 'dibagikan', 'terbagi',
'terbagikan', 'sedia', 'bukan', 'beri', 'nyata', 'tambah', 'buat', 'utama',
'menyediakan', 'oke', 'lebih', 'mampir', 'diberitahu', 'payung', 'sejak', 'manamana',
'kiss', 'suat', 'natuh', 'trida', 'mudah', 'm', 'bakti', 'tatah'

```

```

[]
'sulit','mulai','didapat','dibawah','kebahah','disamakan','dapat','menit','sayang',
'lama','cari','jelas','lewat','objek','bawah','kumpul','rumah','200',
'bagi','sama','semuanya','terakhir','tungguh','menang','kasih',
'bersedia','pohon','masuk','lihat','baru','sebuah',
'000','10','berbagi','momen','terbaru','berguna','berakhir','pendapat',
'mendapatkan','terdapat','didapatkan','bagikan','dibagikan','terbagi',
'terbagikan','sedia','bukan','berl','nyata','tambah','buat','utama',
'menyediakan','oke','lebih','mampir','diberitahu','payung','sejak','manamana',
'kios','awet','patut','tuju','mudah','m','kaki','letak',
'lihat','padan','tepat','jadi','coba','duduk','bawa','tahu',
'capat','bantu','seluruh','tasa',
'ambil','lucu','estetis','karyaan','basir','estetika',
'rekomendasi','santa','benar','kali','cakeppp','cakep','maantaap','nihhhh','karyaanmya','kontemporer',
'menghabiskan','sebuah','pertama','sedang','sepadan','tersedia','disediakan',
'memberikannya','diberikan','memberikan','diberikannya','berikan','diambil','mengambil','diambilkan',
'bagal','berbagal','bersantai','nyantai','asik','asa','amal','cara','lalu','minta','beli','padahal','milih','ajak','buru','cara','tidur','tamu',
'asa','amal','naik','asik','kapan','penting','giveaway','akun','undang','hadiah','uang',
'depan','masalah','heha','hehaforest','sebut','lalu','minta','laku','ayah','suami',
'2','2','pelayanannya','makanah','antarjemput','pelayananannya','pelayanananan','pelayananan','layanananan','layananan',
'saran','maaf','bading','turun','ganti','makin',
'wajib','sekadar','jajah','wan','tim','ganteng','suami',
'bisallahirrahmanirrahim','hehe','yo','positif','senyum','patut',
'milik','ulas','kelling','jaket','nihhhh','maantaap','krumah','celana','pahit','cakeppp','tikung','tangga','lapis',
'masa','pagi','perhati','pribadi','tutup','sepi','sahabat','panggil',
'bethah','asap','belanja','hereng','dewasa','pukul','ha','kelola','kawatir',
'sholatnya','larut','kelompok','beda','keluar','tiba','daerah','tinggal','belakang',
'lungung','selalu','sambut','karyaan','sempurna','bahagia','sejuk','jincha','ij',
'kotor','bersih','na','pak','apa','kalian','kan','dunia','alam','warna','hubung','solusi',
'bertuk','asal','tampak','lalat','rokok','jeep','lava','unik','aman','dingin','luas',
'perempuan','kakak','rene','pacar','wudhu','jarak','batas','penuh',
'pes','sila','aneh','bos','daki','racun','baca','suatu','punya','tak',
'hiruk','pikuk','perosot','top','hati','jika','rindu','cinta','sedih','damai','tanya',
'hehe','semuanya','pantai','bisallah','berapa','buah','hujan','sehat','pandang',
'ceklah','hehe','hebat','ayih','bagus','bahagi','masa','rp','larut','es','belah',
'6','4','8','lambilampu','pedas','besok','sekarang','tahun','dulu','sepat','larjut',
'wajib','sekadar','sekadar','top','spesial','iki','kallurang','sempit','kawatir',
'kunjung','gabung','teman-teman','moga','libur','menang','jalan','stres','pikir',
'mohon','tanggap','cocok','kati','mental','suasana','hadap','sebuah','benarbenar','destinasi',
'ajak','mengajak','indah','selamat','selesai','sore','saudara','terima','untung','admin','ramai','halo','instagram',
'yuk','kompetisi','ikut','minimal','kemari','mari','pekan','sis','jam','sabar','pulih','forest','malam','rasa','cantik',
'kemarin','usia','bu','allah','grup','tani','tenang','romantis','tugas','informasi','umam','hasil','kembang','sebentar',
'desain','beneran','mipip','lila','ehh','ruka','camli','kun','2023','jumat','boger',
'manfaat','ampun','wipai','senyum','pussa','kisah','selesai','gugur','herndatewanya','akar','uit','gerai','energi',
'pakai','batu','kena','rutinitas','diam','bingga','bebas','segala','mulai','siapa','kotak','mini','hitung','halita','hadeuhhh','konsekuensi','tanah','akibat','longsor','nyaw',
'bisnis','capai','amin','anisaakia','batik','aliya711','story','hambat','goy','kahit',
'mantabs','naga','andal','aura','bernadethavania','jas','haru','relatable','auliamh','pipi',
'jadwal','umur','nas','divyuspitarini','nyamankien','wilayah','akhirat','antar',
'serius','loh','uji','berani','perut','35','wahana','nuansa','fittasari','skut','dewiskeraisyah','daryanti0699','bhumi',
'kemah','keri','jaga','tumpuk','paham','beban','cartoonsmithaf','jujur','nilai','salam',
'renervasi','akun','berangkat','kalaham','aksi','bsd','gambar','rahma23','wajibhh','rencana',
'febrina2110','bukankah','nona','musin','kue','dengar','live','camil','interstellar','ompi',
'kolaborasi','warga','tubuh','forester','lelah',
'matahari','ahh','pingin','wefellfilled','ning','piskinurhanifah217','apike','koyone','mengi','total','sesi','sewa','rumput','prowedding',
'jadwal','umur','area','akses','tradisional','cahaya','selfie','gelap','pasta','ji','sapi','hijau','masingmasing',

```

```

[]
'langung','selalu','sambut','karyaan','sempurna','bahagia','sejuk','jincha','ij',
'kotor','bersih','na','pak','apa','kalian','kan','dunia','alam','warna','hubung','solusi',
'bertuk','asal','tampak','lalat','rokok','jeep','lava','unik','aman','dingin','luas',
'perempuan','kakak','rene','pacar','wudhu','jarak','batas','penuh',
'pes','sila','aneh','bos','daki','racun','baca','suatu','punya','tak',
'hiruk','pikuk','perosot','top','hati','jika','rindu','cinta','sedih','damai','tanya',
'hehe','semuanya','pantai','bisallah','berapa','buah','hujan','sehat','pandang',
'ceklah','hehe','hebat','ayih','bagus','bahagi','masa','rp','larut','es','belah',
'6','4','8','lambilampu','pedas','besok','sekarang','tahun','dulu','sepat','larjut',
'wajib','sekadar','sekadar','top','spesial','iki','kallurang','sempit','kawatir',
'kunjung','gabung','teman-teman','moga','libur','menang','jalan','stres','pikir',
'mohon','tanggap','cocok','kati','mental','suasana','hadap','sebuah','benarbenar','destinasi',
'ajak','mengajak','indah','selamat','selesai','sore','saudara','terima','untung','admin','ramai','halo','instagram',
'yuk','kompetisi','ikut','minimal','kemari','mari','pekan','sis','jam','sabar','pulih','forest','malam','rasa','cantik',
'kemarin','usia','bu','allah','grup','tani','tenang','romantis','tugas','informasi','umam','hasil','kembang','sebentar',
'desain','beneran','mipip','lila','ehh','ruka','camli','kun','2023','jumat','boger',
'manfaat','ampun','wipai','senyum','pussa','kisah','selesai','gugur','herndatewanya','akar','uit','gerai','energi',
'pakai','batu','kena','rutinitas','diam','bingga','bebas','segala','mulai','siapa','kotak','mini','hitung','halita','hadeuhhh','konsekuensi','tanah','akibat','longsor','nyaw',
'bisnis','capai','amin','anisaakia','batik','aliya711','story','hambat','goy','kahit',
'mantabs','naga','andal','aura','bernadethavania','jas','haru','relatable','auliamh','pipi',
'jadwal','umur','nas','divyuspitarini','nyamankien','wilayah','akhirat','antar',
'serius','loh','uji','berani','perut','35','wahana','nuansa','fittasari','skut','dewiskeraisyah','daryanti0699','bhumi',
'kemah','keri','jaga','tumpuk','paham','beban','cartoonsmithaf','jujur','nilai','salam',
'renervasi','akun','berangkat','kalaham','aksi','bsd','gambar','rahma23','wajibhh','rencana',
'febrina2110','bukankah','nona','musin','kue','dengar','live','camil','interstellar','ompi',
'kolaborasi','warga','tubuh','forester','lelah',
'matahari','ahh','pingin','wefellfilled','ning','piskinurhanifah217','apike','koyone','mengi','total','sesi','sewa','rumput','prowedding',
'jadwal','umur','area','akses','tradisional','cahaya','selfie','gelap','pasta','ji','sapi','hijau','masingmasing',

```

```

[]
# Buat regex untuk stopwords removal
stopwords_pattern = re.compile(r'\b(' + r'|'.join(re.escape(s) for s in all_stopwords) + r')\b\s*')

def remove_stopwords_from_string(text):
    text = re.sub(stopwords_pattern, '', text)
    text = re.sub(r'(\s+)', ' ', text).strip()
    return text

df_processed['stopword_removed_text'] = df_processed['case_folding'].apply(remove_stopwords_from_string)

# b. Tokenizing
print("Melakukan Tokenizing...")
# Gunakan NLTK word_tokenize
df_processed['tokenizing'] = df_processed['stopword_removed_text'].apply(nltk.word_tokenize)

```

```

df_processed['stopword_removed_text'] = df_processed['case_folding'].apply(remove_stopwords_from_string)

# b. Tokenizing
print("Melakukan Tokenizing...")
# ambilkan nltk word_tokenize
df_processed['tokenizing'] = df_processed['stopword_removed_text'].apply(nltk.word_tokenize)

# c. Filtering (Symbol cleaning on tokens)
print("Melakukan pembersihan simbol dari token (filtering)...")
def clean_single_token(token):
    # Remove any non-alphanumeric characters (like '[ ]', '') from the token
    cleaned_token = re.sub(r'[\s-Za-Z0-9]', '', token)
    return cleaned_token.strip()

def clean_list_of_tokens(token_list):
    if not isinstance(token_list, list):
        return []
    cleaned_list = [clean_single_token(token) for token in token_list]
    # filter out any empty strings that resulted from cleaning
    return [token for token in cleaned_list if token]

df_processed['filtering'] = df_processed['tokenizing'].apply(clean_list_of_tokens)

# d. Stemming
print("Melakukan Stemming...")
factory_stem = StemmerFactory()
stemmer = factory_stem.create_stemmer()

```

```

df_processed['filtering'] = df_processed['tokenizing'].apply(clean_list_of_tokens)

# d. Stemming
print("Melakukan Stemming...")
factory_stem = StemmerFactory()
stemmer = factory_stem.create_stemmer()

def stem_text(tokens):
    if isinstance(tokens, list):
        return [stemmer.stem(word) for word in tokens]
    return [] # Mengembalikan list kosong jika input bukan list

df_processed['stemmed'] = df_processed['filtering'].apply(stem_text)

# e. Kolom 'kalimat' (gabungan token hasil stemming)
print("Membuat kolom 'kalimat'...")
df_processed['kalimat'] = df_processed['stemmed'].apply(lambda x: ' '.join(x) if isinstance(x, list) else '')

print("\nProses preprocessing selesai.")
print("Dataframe hasil preprocessing:")
print(df_processed.head())

# Menampilkan kolom-kolom yang dihasilkan/dibiarkan sesuai permintaan:
print(df_processed.columns.tolist())

```

Requirement already satisfied: Sastrawi in /usr/local/lib/python3.12/dist-packages (1.0.1)
Dataset berhasil dimuat.

```

print("\n--- Verifikasi Stopword --- ")

# Ambil beberapa sampel dari kolom 'kalimat'
sample_kalimat = df_processed['kalimat'].sample(n=10, random_state=42).tolist()

stopwords_found_in_sample = []

for i, text in enumerate(sample_kalimat):
    words = text.split()
    found_in_text = [word for word in words if word in all_stopwords]
    if found_in_text:
        stopwords_found_in_sample.append((i, text, found_in_text))

if stopwords_found_in_sample:
    print("Stopwords ditemukan di beberapa sampel 'kalimat':")
    for idx, text, found_words in stopwords_found_in_sample:
        print(f" Sampel (idx): '{text}'")
        print(f" Stopwords ditemukan: {found_words}")
    else:
        print("Tidak ada stopwords umum yang ditemukan di sampel 'kalimat'.")

print("\n--- Verifikasi selesai --- ")

--- Verifikasi Stopword ---
Stopwords ditemukan di beberapa sampel 'kalimat':
Sampel 1: 'nyaman rekomendasi kunjung'

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ] print("Melakukan stopwords removal ulang pada kolom 'stemmed' dan 'kalimat'...")

# Pastikan stemmer tersedia
if 'stemmer' not in globals():
    from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
    factory_stem = StemmerFactory()
    stemmer = factory_stem.create_stemmer()

# 1. Stem the all_stopwords list for more effective removal on stemmed tokens
stemmed_all_stopwords = [stemmer.stem(word) for word in all_stopwords]
# Remove duplicates and ensure it's a set for faster lookup
stemmed_all_stopwords_set = set(stemmed_all_stopwords)

# 2. Function to remove stopwords from a list of stemmed tokens
def remove_stemmed_stopwords_from_list(stemmed_tokens, list):
    if isinstance(stemmed_tokens, list):
        # Filter out tokens that are in the stemmed_all_stopwords_set and also filter out empty strings
        return [token for token in stemmed_tokens if token not in stemmed_all_stopwords_set and token != '']
    return []

# 3. Apply this new stopwords removal to the 'stemmed' column
df_processed['stemmed'] = df_processed['stemmed'].apply(remove_stemmed_stopwords_from_list)

# 4. Re-create the 'kalimat' column from the updated 'stemmed' column
df_processed['kalimat'] = df_processed['stemmed'].apply(lambda x: ' '.join(x) if isinstance(x, list) else '')

print("Stopword removal ulang pada kolom 'stemmed' dan 'kalimat' telah selesai.")

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ] # 3. Apply this new stopwords removal to the 'stemmed' column
df_processed['stemmed'] = df_processed['stemmed'].apply(remove_stemmed_stopwords_from_list)

# 4. Re-create the 'kalimat' column from the updated 'stemmed' column
df_processed['kalimat'] = df_processed['stemmed'].apply(lambda x: ' '.join(x) if isinstance(x, list) else '')

print("Stopword removal ulang pada kolom 'stemmed' dan 'kalimat' telah selesai.")
print("\n--- Verifikasi Stopword Setelah Pembersihan Ulang ---")

# Ambil beberapa sampel dari kolom 'kalimat' setelah pembersihan ulang
sample_kalimat_recleaned = df_processed['kalimat'].sample(n=10, random_state=42).tolist()

stopwords_found_in_sample_recleaned = []

for i, text in enumerate(sample_kalimat_recleaned):
    words = text.split()
    # Check against the stemmed_all_stopwords_set
    found_in_text = [word for word in words if word in stemmed_all_stopwords_set]
    if found_in_text:
        stopwords_found_in_sample_recleaned.append((i, text, found_in_text))

if stopwords_found_in_sample_recleaned:
    print("Stopwords masih ditemukan di beberapa sampel 'kalimat' setelah pembersihan ulang:")
    for idx, text, found_words in stopwords_found_in_sample_recleaned:
        print(f" Sampel {idx}: '{text}'")
        print(f" Stopwords ditemukan: {found_words}")

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ] print(f" Sampel {idx}: '{text}'")
print(f" Stopwords ditemukan: {found_words}")
else:
    print("Tidak ada stopwords umum (setelah stemming) yang ditemukan di sampel 'kalimat' setelah pembersihan ulang.")

print("\n--- Verifikasi selesai ---")

Melakukan stopwords removal ulang pada kolom 'stemmed' dan 'kalimat'...
Stopword removal ulang pada kolom 'stemmed' dan 'kalimat' telah selesai.

--- Verifikasi Stopword Setelah Pembersihan Ulang ---
Tidak ada stopwords umum (setelah stemming) yang ditemukan di sampel 'kalimat' setelah pembersihan ulang.
--- Verifikasi selesai ---

[ ] output_filename = 'sentiment_maps_update(22).xlsx'
df_processed.to_excel(output_filename, index=False)
print(f"File '{output_filename}' berhasil disimpan dan siap diunduh.")

File 'sentiment_maps_update(21).xlsx' berhasil disimpan dan siap diunduh.

[ ] # Install gensim if not already installed
!pip install gensim

import pandas as pd
from gensim import corpora
from gensim.models.ldamodel import LdaModel

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Q Commands + Code + Text ▶ Run all
Connect

1
# Install gensim if not already installed
!pip install gensim

import pandas as pd
from gensim import corpora
from gensim.models.ldamodel import LdaModel
from gensim.models import CoherenceModel
import numpy as np

# --- Langkah 1: Muat Dataset ---
# Asumsi file 'review_maps_with_sentiment.xlsx' ada di lingkungan colab.
try:
    df = pd.read_excel('sentiment_maps_update(22)_lda_results.xlsx')
    print("Dataset berhasil dimuat.")
except FileNotFoundError:
    print("Error: File 'sentiment_maps_update(22)_lda_results.xlsx' tidak ditemukan. Pastikan file sudah diunggah.")
    # Hentikan eksekusi jika file tidak ditemukan
    exit()
except Exception as e:
    print(f"Terjadi error saat memuat dataset: {e}")
    exit()

# --- Tampilkan semua nama kolom ---
print("Nama-nama kolom yang tersedia di Dataframe Anda:")
print(df.columns)

# --- Langkah 2: Persiapkan Data Teks ---

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Q Commands + Code + Text ▶ Run all
Connect

1
# Menangani nilai yang hilang dan membagi string menjadi daftar token
documents = df[column_name].dropna().apply(lambda x: str(x).split())

if documents.empty:
    print(f"Error: Kolom '{column_name}' kosong atau hanya berisi nilai yang hilang setelah penghapusan NaN.")
    exit()
else:
    print(f"Berhasil memproses kolom '{column_name}'. Jumlah dokumen: (len(documents))")

# Buat kamus (dictionary) dari dokumen tokenized
dictionary = corpora.Dictionary(documents)

# Buat korpus (Bag of Words) dari dokumen
corpus = [dictionary.doc2bow(doc) for doc in documents]
print(f"Kamus dibuat dengan (len(dictionary)) token unik.")
print(f"Korpus dibuat dengan (len(corpus)) dokumen.")

# --- Langkah 3: Lakukan Analisis LDA untuk Rentang Topik 10-26 ---
min_topics = 10
max_topics = 26 # Inklusif
step_topics = 1 # Increment by 1

coherence_scores = []
lda_models = []
num_topics_list = list(range(min_topics, max_topics + 1, step_topics))

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Q Commands + Code + Text ▶ Run all
Connect

1
for num_topics in num_topics_list:
    print(f"Melatih model LDA dengan (num_topics) topik...")
    # Latih model LDA
    # 'auto' untuk alpha dan eta sering memberikan hasil yang baik.
    # num_passes adalah jumlah pass melalui korpus selama pelatihan.
    # iterations adalah jumlah iterasi per dokumen.
    lda_model = LdaModel(corpus=corpus,
                        id2word=dictionary,
                        num_topics=num_topics,
                        random_state=100, # untuk reproduktivitas
                        update_every=1,
                        chunksize=100,
                        passes=10,
                        alpha='auto',
                        eta='auto',
                        per_word_topics=True)
    lda_models.append(lda_model)

    # Hitung Coherence Score (menggunakan 'c_v')
    coherence_model_lda = CoherenceModel(model=lda_model, texts=documents, dictionary=dictionary, coherence='c_v')
    coherence_lda = coherence_model_lda.get_coherence()
    coherence_scores.append(coherence_lda)
    print(f"Skor Koherensi untuk (num_topics) topik: {coherence_lda:.4f}")

# --- Langkah 4: Tampilkan Hasil ---
print("\n--- Skor Koherensi untuk Berbagai Jumlah Topik ---")
for i, num_topics in enumerate(num_topics_list):
    print(f"Jumlah Topik: {num_topics}, Skor Koherensi: {coherence_scores[i]:.4f}")

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ]
best_num_topics = num_topics_list[best_num_topics_idx]
best_lda_model = lda_models[best_num_topics_idx]

print(f"\nJumlah topik optimal berdasarkan skor koherensi: {best_num_topics}")

# Tampilkan kata-kata teratas untuk model terbaik
print(f"\n--- Kata-kata teratas untuk {best_num_topics} Topik (Model Terbaik) ---")
for idx, topic in best_lda_model.print_topics(-1):
    print(f"Topik: {idx} \nkata-kata: {topic}")

# Anda dapat menyimpan model terbaik jika diperlukan:
# best_lda_model.save(f'lda_model_best_{best_num_topics}.gensim')

collecting gensim
Downloading gensim-4.4.0-cp312-cp312-manylinux_2_24_x86_64.manylinux_2_28_x86_64.whl.metadata (8.4 kB)
Requirement already satisfied: numpy>=1.18.5 in /usr/local/lib/python3.12/dist-packages (from gensim) (2.0.2)
Requirement already satisfied: scipy>=1.7.0 in /usr/local/lib/python3.12/dist-packages (from gensim) (1.16.3)
Requirement already satisfied: smart_open>=1.8.1 in /usr/local/lib/python3.12/dist-packages (from gensim) (7.6.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages (from smart_open>=1.8.1->gensim) (2.1.2)
Downloading gensim-4.4.0-cp312-cp312-manylinux_2_24_x86_64.manylinux_2_28_x86_64.whl (27.9 MB)
27.9/27.9 MB 61.0 MB/s eta 0:00:00

Installing collected packages: gensim
Successfully installed gensim-4.4.0
Dataset berhasil dimuat.
Nama-nama kolom yang tersedia di DataFrame Anda:
Index(['title', 'url', 'stars', 'name', 'text', 'english_translation',
       'indonesia_translate', 'sentiment_score', 'sentiment', 'case_folding',
       'stopword_removed_text', 'tokenizing', 'filtering', 'stemmed',
       'kalimat', 'topik', 'dominant_topic_probability', 'aspek'],
      dtype='object', length=16)

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ]
topik_kata = {
    4: "Kata-kata: 0.225*\"menu\" + 0.045*\"hidang\" + 0.034*\"gudeg\" + 0.024*\"sushi\" + 0.018*\"mie\" + 0.017*\"malah\" + 0.016*\"minus\" + 0.013*\"lokal\" + 0.013*\"barat\" + 0.010*\"matah\"",
    5: "Topik: 5",
    6: "Kata-kata: 0.195*\"main\" + 0.164*\"taman\" + 0.049*\"motor\" + 0.040*\"fasilitas\" + 0.018*\"jenis\" + 0.016*\"arah\" + 0.014*\"inti\" + 0.013*\"lingkung\" + 0.013*\"alpuat\" + 0.011*\"takj\"",
    7: "Topik: 6",
    8: "Kata-kata: 0.474*\"foto\" + 0.199*\"spot\" + 0.080*\"tarik\" + 0.022*\"alami\" + 0.013*\"bilang\" + 0.009*\"latar\" + 0.007*\"lepas\" + 0.006*\"kelas\" + 0.005*\"tata\" + 0.005*\"kaferestoran\"",
    9: "Topik: 7",
    10: "Kata-kata: 0.236*\"keluarga\" + 0.189*\"wisata\" + 0.185*\"hutan\" + 0.091*\"minum\" + 0.028*\"kualitas\" + 0.021*\"relaksasi\" + 0.017*\"estetik\" + 0.015*\"rapi\" + 0.015*\"rekreasi\" + 0.014*\"kebudayaan\"",
    11: "Kata-kata: 0.385*\"ramah\" + 0.096*\"staf\" + 0.031*\"jual\" + 0.027*\"dapur\" + 0.014*\"henti\" + 0.013*\"plinggir\" + 0.013*\"profesional\" + 0.013*\"cuma\" + 0.012*\"sopan\" + 0.010*\"kebudayaan\"",
    12: "Topik: 8"
}

import matplotlib.pyplot as plt

# --- Langkah 5: Visualisasikan Skor Koherensi ---
plt.figure(figsize=(10, 6))
plt.plot(num_topics_list, coherence_scores, marker='o')
plt.xlabel("Jumlah Topik")
plt.ylabel("Skor Koherensi (c.v)")
plt.title("Skor Koherensi Model LDA berdasarkan Jumlah Topik")
plt.xticks(num_topics_list) # Menampilkan semua label sumbu x
plt.grid(True)

# Tandai titik optimal
plt.axvline(x=best_num_topics, color='r', linestyle='--', label=f'Jumlah Topik Optimal: {best_num_topics}')
plt.legend()
plt.show()

```

```

Untitled11.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Connect

[ ]
from pandas import DataFrame

# Ekstrak kata-kata teratas dan bobotnya untuk setiap topik dari best_lda_model
topic_data = []
for idx, topic in best_lda_model.print_topics(-1):
    topic_words = [word.split(":") for word in topic.split(": ")]
    # Memformat ulang agar lebih mudah dibaca
    formatted_words = []
    for prob, word in topic_words:
        # Menghapus tanda kutip dari kata
        word = word.replace('"', '')
        formatted_words.append(f"{word} ({float(prob):.3f})")
    topic_data.append([f"Topik {idx}", " ".join(formatted_words)])

# Buat DataFrame
topics_df_optimal = DataFrame(topic_data, columns=['Topik', 'Top Words (Probability)'])

# Simpan DataFrame ke file Excel dengan nama yang mencerminkan jumlah topik optimal
output_excel_file_optimal = f'lda_top_words_{best_num_topics}_optimal_topics.xlsx'
topics_df_optimal.to_excel(output_excel_file_optimal, index=False)

print(f"File Excel '{output_excel_file_optimal}' berhasil dibuat. Anda dapat mengunduhnya dari panel file Colab.")

File Excel 'lda_top_words_10_optimal_topics.xlsx' berhasil dibuat. Anda dapat mengunduhnya dari panel file Colab.

```