Journal Software, Hardware and Information Technology | SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

Identifikasi Daya Tarik Wisata Benteng Rotterdam Berdasarkan Ulasan Wisatawan Menggunakan *Topic Modeling*

Izmy Alwiah Musdar [1*], Husni Angriani [2], St Muriati [3]
[1] Program Studi Sistem Informasi, Universitas Islam Negeri Alauddin Makassar, Indonesia
[2] Program Studi Sistem Informasi, STMIK KHARISMA Makassar, Indonesia
[3] Program Studi Pendidikan Ilmu Pengetahuan Alam, Universitas Bosowa, Indonesia
[1] izmy.alwiah@uin-alauddin.ac.id, [2] husniangriani@kharisma.ac.id, [3] st.muriati@universitasbosowa.ac.id

Diajukan: 27/05/2026 | Direvisi: 20/6/2026 | Diterima: 22/06/2026 | Diterbitkan: 30/06/2026

Abstrak

Perkembangan platform ulasan daring memberikan peluang untuk menggali informasi mengenai persepsi wisatawan terhadap suatu destinasi wisata. Penelitian ini bertujuan untuk mengidentifikasi daya tarik utama Benteng Rotterdam berdasarkan ulasan wisatawan menggunakan metode topic modeling Latent Dirichlet Allocation (LDA). Data penelitian diperoleh dari TripAdvisor melalui teknik *web scraping* menggunakan *library* Selenium dan BeautifulSoup. Proses *scraping* data dilakukan pada Desember 2023 yang menghasilkan sebanyak 679 ulasan berbahasa Indonesia. Ulasan yang telah dikumpulkan kemudian diproses melalui tahapan *preprocessing* yang meliputi *cleaning*, *tokenization*, *stemming*, dan *stopword removal*. Pemodelan topik dilakukan menggunakan *library Gensim* pada bahasa pemrograman Python, sedangkan penentuan jumlah topik optimal dilakukan menggunakan *coherence score*. Visualisasi menggunakan *word cloud* dan *intertopic distance map* digunakan untuk membantu dalam interpretasi hubungan antar topik. Hasil penelitian menunjukkan bahwa model LDA berhasil menghasilkan enam topik dengan nilai *coherence score* sebesar 0,397836 yang, berdasarkan interpretasi kata dominan dan hasil visualisasi, dapat diidentifikasi sebagai daya tarik wisata Benteng Rotterdam, meliputi nilai sejarah bangunan, suasana dan kondisi bangunan, lokasi strategis, pengalaman sosial dan budaya, fasilitas pendukung, serta biaya kunjungan yang terjangkau. Temuan penelitian ini dapat dimanfaatkan sebagai masukan bagi pengelola wisata dalam menyusun strategi promosi dan pengembangan layanan berbasis ulasan wisatawan.

Kata Kunci: LDA, Pariwisata, *Topic Modeling*, Ulasan Wisatawan

Abstract

The development of online review platforms provides opportunities to explore information regarding tourists' perceptions of tourism destinations. This study aims to identify the main attractions of Benteng Rotterdam based on tourist reviews using the Latent Dirichlet Allocation (LDA) topic modeling method. The research data were obtained from TripAdvisor through a web scraping technique using Selenium and BeautifulSoup libraries. Data scraping was performed in December 2023, resulting in the collection of 679 reviews written in Indonesian. The collected reviews were then subjected to a series of preprocessing steps, namely cleaning, tokenization, stemming, and stopword removal. Topic modeling was conducted using the Gensim library in the Python programming language, while the optimal number of topics was determined using the coherence score. Visualization using word clouds and intertopic distance maps was applied to assist in interpreting the relationships among topics. The results show that the LDA model successfully generated six topics which, based on the interpretation of dominant keywords and visualization results, can be identified as the main tourist attractions of Benteng Rotterdam, including the historical value of the building, building atmosphere and condition, strategic location, social and cultural experiences, and affordable entrance fees. The findings of this study can serve as input for tourism managers in developing promotional strategies and improving tourism services based on visitor reviews.

Keywords**:** *LDA, Topic Modeling, Tourism, Tourist Reviews*

This work is an open access article and licensed under a Creative Commons Attribution-NonCommercial ShareAlike 4.0 International (CC BY-NC-SA 4.0). Copyright (C) Authors.

1. PENDAHULUAN

Industri pariwisata berkembang pesat seiring kemudahan akses informasi dan transportasi. Calon wisatawan kini cenderung mencari referensi secara daring—baik untuk menentukan tujuan maupun untuk memperkirakan kualitas layanan yang akan diterima [1],[2] . Salah satu sumber informasi paling berpengaruh ialah ulasan daring (*electronic word-of-mouth)*

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

yang ditulis oleh wisatawan pasca kunjungan [3], [4], [5], [6] Penelitian terdahulu menunjukkan bahwa ulasan tersebut secara signifikan memengaruhi keputusan berkunjung, menunjukkan daya tarik visual destinasi wisata, serta kepuasan wisatawan [7]. Oleh karena itu, memahami isi dan topik-topik dominan dalam ulasan menjadi penting bagi pengelola destinasi untuk menyusun strategi promosi, peningkatan layanan, dan pengembangan produk wisata [8].

Salah satu cabang NLP yang sesuai untuk menggali tema besar dalam kumpulan ulasan ialah *topic modeling* [9]. *Topic modeling* adalah teknik *unsupervised machine learning* yang secara otomatis menemukan sekumpulan topik berdasarkan pola kemunculan bersama (*Co-occurrence*) kata dalam korpus dokumen [8], [10]. Setiap topik direpresentasikan sebagai distribusi probabilitas atas kosakata dimana kata-kata dengan probabilitas tertinggi secara intuitif mencerminkan tema umum yang dibicarakan [11]. Beberapa algoritma *topic modeling* yang umum digunakan meliputi Latent Semantic Indexing (LSI), Probabilistic Latent Semantic Indexing (PLSI), Latent Dirichlet Allocation (LDA), dan Correlated Topic Model (CTM) [12]. Pada penelitian ini, Latent Dirichlet Allocation (LDA) dipilih karena Algoritma LDA memiliki beberapa keunggulan, antara lain mudah diimplementasikan untuk menemukan topik utama, kemampuan memproses data teks dengan cepat [13], hasil pemodelan yang stabil selama proses iterasi, kemampuan yang baik dalam menangani data teks yang beragam dan mengandung *noise*, serta didukung oleh implementasi yang efisien pada berbagai *library* seperti Gensim dan MALLET [12], [14]. Selain itu, penelitian ini berfokus pada identifikasi tema dan daya tarik wisata yang mudah diinterpretasikan. Oleh karena itu, LDA dipilih karena menghasilkan representasi topik berupa distribusi probabilitas kata yang relatif mudah dipahami dan dijelaskan.

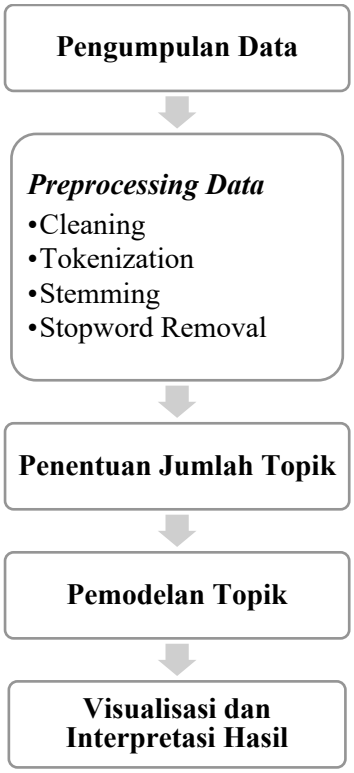
Penelitian ini mengaplikasikan LDA untuk memodelkan topik-topik yang muncul dalam ulasan berbahasa Indonesia Benteng Rotterdam, sebuah objek wisata sejarah terkemuka yang terletak di pusat Kota Makassar. Meskipun Benteng Rotterdam dikenal sebagai destinasi wisata yang populer, studi mengenai pengalaman wisatawan di objek wisata ini masih didominasi oleh pendekatan kualitatif konvensional, seperti observasi dan wawancara [15], [16], [17], [18]. Pendekatan tersebut memiliki keterbatasan dalam menangkap opini wisatawan dalam jumlah besar dibandingkan analisis berbasis ulasan daring dan teknik *Natural Language Processing* (NLP). Selain itu, ulasan wisata sejarah umumnya memuat beragam aspek secara bersamaan, seperti nilai sejarah, budaya, kondisi lingkungan, fasilitas, dan pengalaman pribadi wisatawan, sehingga menarik untuk dianalisis menggunakan *topic modeling*. Dengan mengetahui topik-topik dominan yang dibicarakan wisatawan, pengelola dapat menyusun narasi promosi yang selaras dengan harapan pasar sekaligus merencanakan perbaikan berdasarkan masukan wisatawan.

Dengan latar belakang tersebut, penelitian ini bertujuan untuk menghasilkan model LDA yang digunakan untuk menemukan topik-topik utama pada ulasan objek wisata Benteng Rotterdam, menginterpretasikan setiap topik melalui kata kunci dan visualisasi *word cloud*, serta merumuskan daya tarik utama yang layak dipromosikan agar kunjungan ke objek wisata ini semakin meningkat.

2. METODE PENELITIAN

Pada bagian ini dijelaskan tahapan penelitian yang dilakukan untuk membangun model *topic modeling* pada ulasan wisatawan Benteng Rotterdam. Tahapan penelitian meliputi pengumpulan data, *preprocessing data*, pembentukan *term dictionary* dan *corpus*, penentuan jumlah topik optimal, pemodelan topik menggunakan Latent Dirichlet Allocation (LDA), serta visualisasi hasil pemodelan. Tahapan penelitian ini ditunjukkan pada Gambar 1.

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa ulasan pengunjung objek wisata Benteng Rotterdam yang diperoleh dari platform TripAdvisor Indonesia. Ulasan yang digunakan dibatasi hanya pada ulasan berbahasa Indonesia untuk menjaga konsistensi bahasa pada proses analisis teks. Proses pengumpulan data dilakukan menggunakan teknik web scraping dengan memanfaatkan library Selenium dan BeautifulSoup pada bahasa pemrograman Python. Scraping data dilakukan pada Bulan Desember 2023. Data hasil *scraping* kemudian disimpan dalam format .csv untuk memudahkan proses pengolahan lebih lanjut.

2.2. *Preprocessing Data*

Tahap awal pengolahan data dilakukan melalui *preprocessing* untuk membersihkan dan menyiapkan data teks sebelum proses pemodelan topik. Tahapan *preprocessing* yang dilakukan meliputi:

- *Cleaning*, yaitu penghapusan simbol, tanda baca, angka, tautan (*URL*), dan karakter khusus lainnya;
- *Tokenization*, yaitu pemecahan kalimat menjadi kumpulan kata atau token;
- *Stemming*, yaitu mengubah kata menjadi bentuk kata dasar.
- *Stopword removal*, yaitu menghapus kata-kata umum yang tidak memiliki makna signifikan dalam pembentukan topik.

Hasil *preprocessing* kemudian digunakan untuk membentuk *term dictionary* menggunakan *library* Gensim [19]. *Term dictionary* berisi daftar seluruh kata unik yang muncul dalam kumpulan dokumen. Selanjutnya dilakukan pembentukan representasi dokumen menggunakan metode *Bag of Words* (BoW) untuk menghitung frekuensi kemunculan setiap kata

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

pada masing-masing dokumen. Beberapa *library* yang digunakan dalam tahapan ini adalah re [20], nltk [21] dan sastrawi [22].

2.3. Penentuan Jumlah Topik

Salah satu tantangan dalam topic modeling adalah proses menentukan jumlah topik yang tepat [23]. Pendekatan yang dapat digunakan untuk menentukan jumlah topik optimal adalah penggunaan nilai *coherence score*. *Coherence score* digunakan untuk mengukur tingkat keterkaitan semantik antar kata dalam suatu topik, sehingga dapat digunakan untuk mengevaluasi kualitas topik yang dihasilkan model. Formula untuk menghitung *coherence score* dapat dilihat pada persamaan 1 berikut : [24]

$$coherence\ (V) = \ score\ (v_i, v_j, \in) = \log \frac{D(v_i, v_j)+\in}{D(v_j)} \qquad (1)$$

dimana V merupakan kumpulan kata dalam sebuah topik, $\in$ adalah variabel untuk menjamin nilai tetap positif (umumnya bernilai 1), $D(v_i, v_j)$ menunjukkan jumlah dokumen yang berisi kata v_i dan v_j, $D(v_j)$ berarti jumlah dokumen dengan kata v_j, $Score\ (v_i, v_j, \in)$ menunjukkan *coherence score* untuk kata v_i, v_j, dan $coherence\ (V)$ adalah jumlah tingkat kemiripan antara kata-kata yang ada di dalam topik V.

2.4. Pemodelan Topik

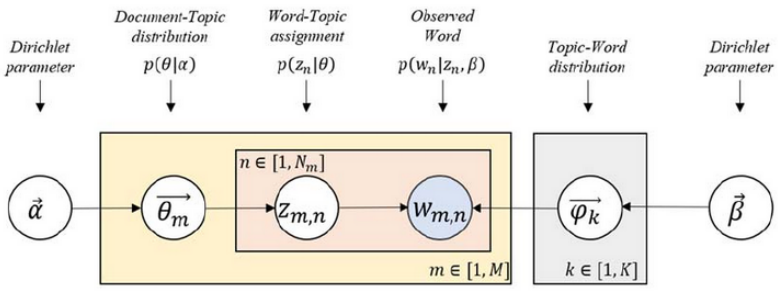
Pemodelan topik dilakukan menggunakan algoritma *Latent Dirichlet Allocation* (LDA). Implementasi algoritma dilakukan menggunakan library Gensim [19] pada bahasa pemrograman Python. Tahapan pemodelan topik pada penelitian ini terdiri dari pembentukan *term dictionary* dan pemodelan menggunakan LDA.

Pembentukan *Term Dictionary* dan *Corpus*

Data ulasan yang telah melalui tahap *preprocessing* selanjutnya digunakan untuk membentuk *term dictionary* menggunakan *library* Gensim. Term dictionary berisi daftar seluruh kata unik yang muncul pada kumpulan dokumen. Setelah itu dilakukan pembentukan corpus menggunakan metode *Bag of Words* (BoW), yaitu representasi dokumen dalam bentuk pasangan kata dan frekuensi kemunculannya pada setiap dokumen. Representasi ini digunakan sebagai masukan dalam proses pemodelan topik menggunakan algoritma LDA.

Pemodelan Topik Menggunakan LDA

Pemodelan topik dilakukan menggunakan algoritma Latent Dirichlet Allocation (LDA). Algoritma LDA digunakan untuk menemukan topik-topik tersembunyi berdasarkan pola kemunculan kata dalam kumpulan dokumen. Setiap topik direpresentasikan sebagai distribusi probabilitas kata, sedangkan setiap dokumen direpresentasikan sebagai distribusi probabilitas topik. Hasil dari proses ini berupa sekumpulan topik beserta kata-kata dominan yang merepresentasikan tema utama pada ulasan wisatawan.



Gambar 2. Diagram Model LDA [25]

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

Secara umum pemodelan menggunakan LDA divisualisasikan pada Gambar 2. setiap dokumen direpresentasikan sebagai distribusi probabilitas topik (θ), sedangkan setiap topik direpresentasikan sebagai distribusi probabilitas kata (φ). Untuk setiap kata dalam dokumen, LDA memilih topik tertentu (z) berdasarkan distribusi topik dokumen, kemudian menghasilkan kata (w) berdasarkan distribusi kata pada topik tersebut. Parameter α (alpha) digunakan untuk mengatur distribusi topik pada dokumen, sedangkan β (beta) digunakan untuk mengatur distribusi kata pada topik. Melalui proses ini, LDA dapat menemukan topik-topik tersembunyi berdasarkan pola kemunculan kata dalam kumpulan dokumen. Implementasi LDA pada penelitian ini memanfaatkan *library* gensim dengan pembuatan model menggunakan parameter random_state=100, chunksize=10, passes=10, alpha='symmetric', dan iterations=100.

2.5. Visualisasi dan Interpretasi Hasil

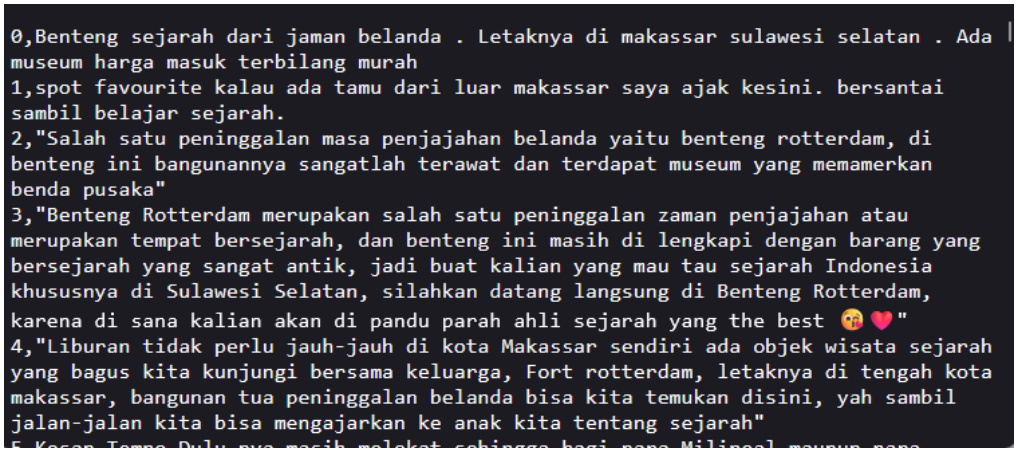
Hasil pemodelan topik berupa sekumpulan kata dengan probabilitas tertentu pada masing-masing topik. Karena topik yang dihasilkan oleh algoritma LDA direpresentasikan dalam bentuk distribusi kata, diperlukan visualisasi untuk mempermudah interpretasi dan identifikasi tema utama dari setiap topik. Pada penelitian ini, visualisasi dilakukan menggunakan *word cloud* untuk menampilkan kata-kata dominan pada setiap topik berdasarkan tingkat kemunculannya. Selain itu, visualisasi juga dilakukan menggunakan *Intertopic Distance Map* untuk menggambarkan hubungan dan jarak antar topik yang dihasilkan model LDA. Visualisasi ini membantu dalam memahami tingkat kedekatan maupun perbedaan antar topik yang terbentuk dari kumpulan ulasan wisatawan.

3. HASIL PENELITIAN DAN PEMBAHASAN

Pada bagian ini disajikan hasil *topic modeling* menggunakan metode LDA terhadap ulasan wisatawan Benteng Rotterdam. Hasil pemodelan digunakan untuk mengidentifikasi topik-topik utama yang merepresentasikan daya tarik wisata berdasarkan persepsi pengunjung pada ulasan daring. Selain itu, dilakukan interpretasi terhadap setiap topik melalui kata-kata dominan dan visualisasi topik untuk memperoleh pemahaman mengenai aspek-aspek yang paling banyak dibicarakan wisatawan.

3.1. Dataset

Dataset di penelitian ini diperoleh dengan melakukan scraping ulasan objek wisata Benteng Rotterdam di Tripadvisor. Berdasarkan hasil *scraping*, total ulasan yang diperoleh sebanyak 679 ulasan wisatawan. Data ulasan ini disimpan dalam file reviewRotterdam.csv. Potongan tangkapan isi layar file tersebut ditunjukkan pada Gambar 3.



Gambar 3. Hasil *Scraping* Data Ulasan

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

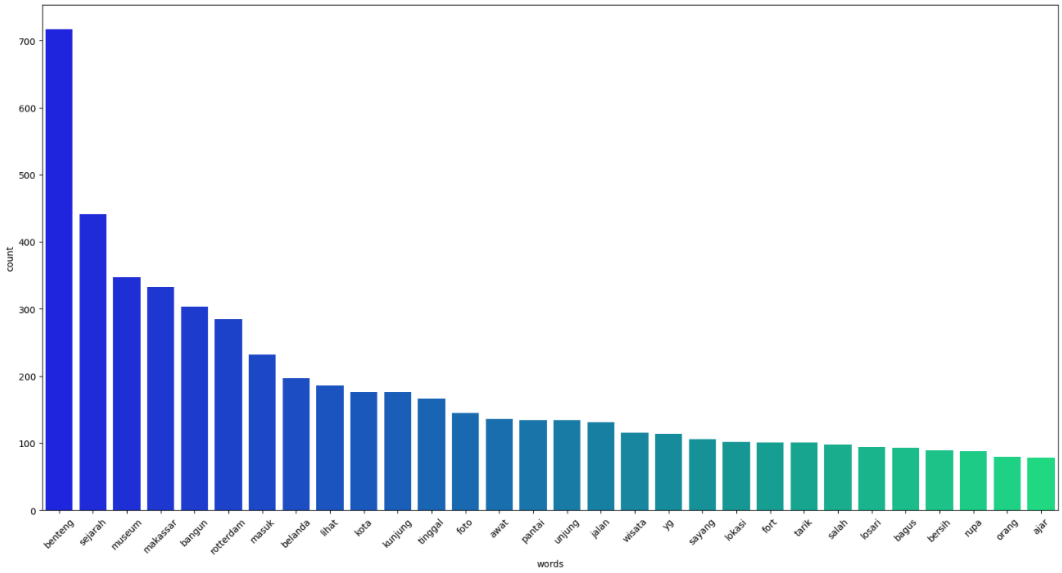
3.2. Hasil *Preprocessing Data*

Hasil *preprocessing* data disimpan dalam file bernama out.csv. File ini berisi 4 kolom data. Kolom pertama adalah indeks, kolom kedua berisi ulasan sebelum melewati *preprocessing*, kolom ketiga berisi hasil *preprocessing* yang dikonversi menjadi *list* sehingga satu satu elemen list menunjukkan satu *token*/kata dalam ulasan, dan kolom keempat hasil *preprocessing* berbentuk teks ulasan utuh. Potongan tangkapan layar file out.csv ditunjukkan pada Gambar 4.

```
,Review_Text,Review_Clean_List,Review_Clean
0,Benteng sejarah dari jaman belanda . Letaknya di makassar sulawesi selatan
. Ada museum harga masuk terbilang murah,"['benteng', 'sejarah', 'jaman',
'belanda', 'letak', 'makassar', 'sulawesi', 'selatan', 'museum', 'harga',
'masuk', 'bilang', 'murah']",benteng sejarah jaman belanda letak makassar
sulawesi selatan museum harga masuk bilang murah
1,spot favourite kalau ada tamu dari luar makassar saya ajak kesini.
bersantai sambil belajar sejarah.,"['spot', 'favourite', 'tamu', 'makassar',
'ajak', 'kesini', 'santa', 'ajar', 'sejarah']",spot favourite tamu makassar
ajak kesini santa ajar sejarah
2,"Salah satu peninggalan masa penjajahan belanda yaitu benteng rotterdam,
di benteng ini bangunannya sangatlah terawat dan terdapat museum yang
memamerkan benda pusaka","['salah', 'tinggal', 'jajah', 'belanda',
'benteng', 'rotterdam', 'benteng', 'bangun', 'awat', 'museum', 'pamer',
'benda', 'pusaka']",salah tinggal jajah belanda benteng rotterdam benteng
bangun awat museum pamer benda pusaka
```

Gambar 4. Tangkapan Layar *File* out.csv

Pada tahapan *stopword removal*, kata-kata umum yang dihilangkan tidak hanya menggunakan daftar kata *stopwords libarary* Sastrawi, tetapi juga mengilangkan beberapa kata dengan frekuensi kemunculan tertinggi yang tidak memiliki makna yang merujuk spesifik ke topik tertentu. Kata-kata ini diidentifikasi melalui pembuatan grafik tiga puluh kata dengan frekuensi kemunculan tertinggi yang ditunjukkan pada Gambar 5. Berdasarkan grafik tersebut kata lainnya yang dianggap sebagai *stopwords* adalah ‘benteng’, ‘rotterdam’, ‘makassar’, ‘nya’, ‘yg’. Hasil penghilangan *stopwords* tersebut ditunjukkan dengan *word cloud* Gambar 6.



Gambar 5. Grafik Batang Frekuensi Kemunculan Kata Tertinggi

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

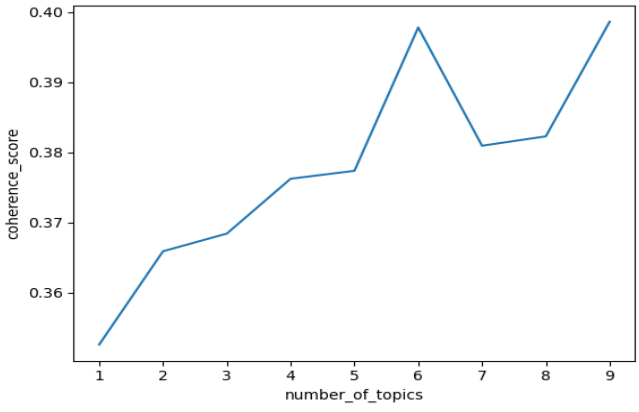
Gambar 6. *Word Cloud* Kata Setelah *Preprocessing*

3.3. Hasil Penentuan Jumlah Topik Optimal

Pada penelitian ini, jumlah topik optimal ditentukan dengan menghitung *coherence score* untuk topik berjumlah 1 hingga 9 menggunakan *library* gensim. Hasil *coherence score* ditunjukkan pada Tabel 1 dan divisualisasikan menggunakan diagram garis pada Gambar 7.

Tabel 1. Nilai *Coherence Score*

Jumlah Topik	Coherence Score
1	0.352593
2	0.365903
3	0.368420
4	0.376248
5	0.377378
6	0.397836
7	0.380960
8	0.382296
9	0.398650



Gambar 7. Diagram Garis Perubahan *Coherence Score* Terhadap Jumlah Topik

Berdasarkan hasil visualisasi pada Gambar 7, nilai *coherence score* mengalami peningkatan hingga mencapai 0,397836 pada jumlah topik ke-6, kemudian mengalami penurunan pada topik ke-7. Meskipun nilai *coherence score* tertinggi diperoleh pada sembilan topik sebesar

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

0,398650, selisihnya dengan model enam topik hanya sebesar 0,000814. Peningkatan yang sangat kecil tersebut menunjukkan bahwa penambahan jumlah topik setelah enam topik tidak memberikan peningkatan kualitas model yang signifikan. Oleh karena itu, penelitian ini memilih enam topik sebagai jumlah topik optimal karena telah memberikan kualitas topik yang baik dengan tingkat interpretasi yang lebih sederhana dibandingkan model dengan jumlah topik yang lebih banyak.

3.4. Hasil Ekstraksi Topik

Hasil pemodelan menggunakan algoritma Latent Dirichlet Allocation (LDA) menghasilkan enam topik utama dari ulasan wisatawan Benteng Rotterdam. Setiap topik direpresentasikan oleh sekumpulan kata dengan probabilitas kemunculan tertinggi yang mencerminkan tema pembahasan pada ulasan pengguna. Hasil ekstraksi topik ditunjukkan pada Tabel 2.

Tabel 2. Hasil Ekstraksi Topik

	Topik 1
Interpretasi Topik	Lokasi dan fasilitas di sekitar objek wisata Benteng Rotterdam
Distribusi Kata	0.074*"hotel" + 0.050*"pantai" + 0.043*"jam" + 0.041*"sisi" + 0.032*"losari" + 0.027*"pusat" + 0.020*"malam" + 0.019*"bugis" + 0.018*"toko" + 0.017*"kaki"
Contoh Ulasan	"Bagus, view sangat-sangat bagus, pusat kota, dekat dengan pantai Losari, dekat dengan Hotel Arthama, transportasi lancar, dekat dengan pusat perbelanjaan" "rekomendasi wisata sejarah....berdekatan dengan pantai losari, dan daerah keramaian. banyak kunjungan sekolah dan pendidikan ke tempat ini, benteng yang berbentuk penyu"
	Topik 2
Interpretasi Topik	Tempat belajar bahasa Inggris
Distribusi Kata	0.078*"bahasa" + 0.071*"inggris" + 0.032*"dinding" + 0.026*"menit" + 0.026*"pikir" + 0.024*"ambil" + 0.019*"coba" + 0.016*"alam" + 0.015*"kamar" + 0.015*"bicara"
Contoh Ulasan	"Lokasinya tidak jauh dari pantai Losari. Dekat dengan pusat kuliner seafood dan tempat hiburan lainnya. Uniknya, museum ini tetap Kokok dengan bangunan yang ada sejak jaman dulu. Hanya saja tetap terawat dan terjaga kebersihannya. Jika berkunjung ke sana, kita juga akan menemui beberapa kelompok kecil remaja yang tengah asik belajar bahasa Inggris dengan komunitas nya, sharing ilmu sambil sesekali bercanda, bermain,menggunakan bahasa Inggris sederhana. Dan pastinya gratis. Mereka sering berkumpul di bawah pohon atau di salah satu teras, yang memang diijinkan oleh pihak pengelola. Sesuatu yang positif untuk anak-anak sekolah. Belajar tidak harus serius kan..Semangat anak Indonesia dan salut untuk pengelola"
	Topik 3
Interpretasi Topik	Suasana dan kondisi bangunan
Distribusi Kata	0.074*"unjung" + 0.058*"sejarah" + 0.051*"tarik" + 0.043*"kunjung" + 0.035*"museum" + 0.029*"pergi" + 0.028*"bersih" + 0.028*"foto" + 0.026*"awat" + 0.024*"bangun"
Contoh Ulasan	"Salah satu ikon sejarah kota Makassar adalah Benteng Rotterdam. Kini, Benteng Rotterdam menjadi salah satu destinasi wisata paling popular di Makassar, baik bagi wisatawan asing maupun wisatawan lokal karena banyak spot spot kece buat foto foto selain itu di Benteng Rotterdam juga ada museum yang bisa disinggahi.Untuk kalian yang sedang berkunjung ke Makassar jangan lupa mampir ya…"

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

	"Salah satu peninggalan masa penjajahan belanda yaitu benteng rotterdam, di benteng ini bangunannya sangatlah terawat dan terdapat museum yang memamerkan benda pusaka"
	Topik 4
Interpretasi Topik	Museum peninggalan Belanda dan tempat penahanan Pangeran Diponegoro
Distribusi Kata	0.104*"lihat" + 0.050*"layak" + 0.048*"buah" + 0.026*"bangun" + 0.023*"museum" + 0.022*"belanda" + 0.021*"kolonial" + 0.018*"tampil" + 0.017*"ambil" + 0.014*"diponegoro"
Contoh Ulasan	"Liburan tidak perlu jauh-jauh di kota Makassar sendiri ada objek wisata Sejarah yang bagus kita kunjungi bersama keluarga, Fort rotterdam, letaknya di Tengah kota makassar, bangunan tua peninggalan belanda bisa kita temukan disini, yah sambil jalan-jalan kita bisa mengajarkan ke anak kita tentang sejarah" "Bangunan ini sejatinya bangunan yang memiliki nilai sejarah yang sangat tinggi. Kita bisa belajar bagaimana perjuangan Sultan Hasanuddin dalam melawan penjajahan belanda, dan kita juga bisa belajar tentang kisah pangeran diponegoro yang dipenjara oleh belanda dalam benteng ini. Sy pribadi sangat mengagumi dan menaruh rasa hormat yang besar terhadap kedua tokoh pahlawan bangsa tersebut. Namun sayang memang, penghargaan terhadap tempat bersejarah ini sangatlah rendah menurut saya. Adanya coretan dan pemeliharaan museum serta ruang bersejarah lainnya masih belum bisa memuaskan saya…."
	Topik 5
Interpretasi Topik	Bangunan bersejarah
Distribusi Kata	0.047*"sejarah" + 0.044*"jalan" + 0.041*"orang" + 0.037*"fort" + 0.037*"milik" + 0.031*"museum" + 0.031*"bangun" + 0.031*"bagus" + 0.026*"temu" + 0.021*"kota"
Contoh Ulasan	"spot favourite kalau ada tamu dari luar makassar saya ajak kesini. Bersantai sambil belajar sejarah" "Benteng Rotterdam merupakan salah satu peninggalan zaman penjajahan atau merupakan tempat bersejarah, dan benteng ini masih di lengkapi dengan barang yang bersejarah yang sangat antik, jadi buat kalian yang mau tau Sejarah Indonesia khususnya di Sulawesi Selatan, silahkan datang langsung di Benteng Rotterdam, karena di sana kalian akan di pandu parah ahli sejarah yang the best"
	Topik 6
Interpretasi Topik	Biaya masuk
Distribusi Kata	0.089*"masuk" + 0.067*"museum" + 0.043*"biaya" + 0.034*"bayar" + 0.027*"tiket" + 0.024*"galigo" + 0.023*"nikmat" + 0.021*"la" + 0.018*"rp" + 0.018*"pintu"
Contoh Ulasan	"Benteng sejarah dari jaman belanda . Letaknya di makassar sulawesi selatan . Ada museum harga masuk terbilang murah" "Saya mengunjungi salah satu situs peninggalan Belanda ini saat ada konser. Meski ada persiapan konser untunglah tetap buka. Arsitektur nya menarik sama seperti peninggalan Belanda di Indonesia pada umumnya. Ada museum nya juga dan hanya membayar 5000 rupiah untuk masuk. Spot menarik untuk berfoto juga banyak"

3.5. Visualisasi Hasil Pemodelan dan Interpretasi Topik

Hasil ekstraksi topik selanjutnya divisualisasikan menggunakan word cloud dan intertopic distance map. Visualisasi *word cloud* digunakan untuk mempermudah interpretasi

Journal Software, Hardware and Information Technology | SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

setiap topik berdasarkan dominasi kata-kata yang muncul pada masing-masing topik. Semakin besar ukuran kata pada visualisasi, maka semakin tinggi frekuensi kemunculan kata tersebut dalam topik terkait. Sementara itu, visualisasi *intertopic distance map* digunakan untuk melihat hubungan dan jarak antar topik yang dihasilkan model LDA. Visualisasi ini membantu dalam memahami tingkat kemiripan maupun perbedaan antar topik berdasarkan distribusi kata penyusunnya. Topik yang berdekatan memiliki distribusi kata yang lebih mirip dibandingkan topik yang berjauhan. Visualisasi *word cloud* untuk setiap topik ditunjukkan pada Gambar 8, sedangkan visualisasi *intertopic distance map* ditunjukkan pada Gambar 9.

Gambar 8. *Word cloud* Hasil Ekstraksi Topik

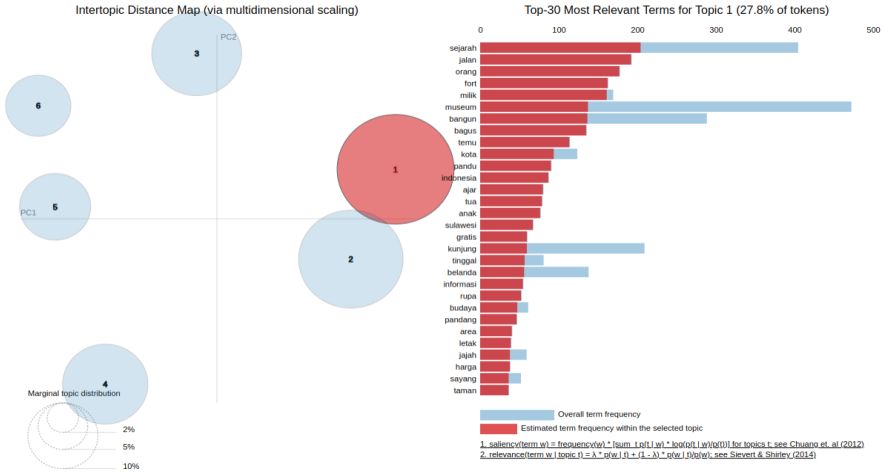
Berdasarkan hasil ekstraksi topik yang disajikan dalam *word cloud* pada Gambar 8, diperoleh beberapa *insight* mengenai persepsi dan perhatian utama wisatawan terhadap Benteng Rotterdam, yaitu :

- Topik pertama berkaitan dengan Lokasi Benteng Rotterdam dan fasilitas di sekitarnya, seperti hotel, Pantai Losari, pusat oleh-oleh, dan area kota. Hal ini menunjukkan bahwa wisatawan tidak hanya membahas objek wisata utama, tetapi juga mempertimbangkan aksesibilitas dan keberadaan fasilitas pendukung di sekitar lokasi wisata.
- Topik kedua menunjukkan bahwa Benteng Rotterdam dimanfaatkan sebagai tempat untuk belajar dan berlatih bahasa Inggris melalui interaksi dengan wisatawan asing. Temuan ini

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

mengindikasikan bahwa kawasan wisata juga memiliki fungsi sosial dan edukatif bagi masyarakat maupun pengunjung lokal.

- Topik ketiga menggambarkan persepsi wisatawan terhadap kondisi fisik Benteng Rotterdam sebagai museum dan situs sejarah. Kata-kata seperti "bersih", "terawat", "indah", dan "foto" menunjukkan bahwa suasana dan kondisi bangunan menjadi faktor penting yang meningkatkan daya tarik wisatawan, khususnya untuk aktivitas fotografi.
- Topik keempat berkaitan dengan nilai historis Benteng Rotterdam sebagai peninggalan kolonial Belanda dan tempat penahanan Pangeran Diponegoro. Topik ini menunjukkan bahwa aspek sejarah menjadi salah satu daya tarik utama yang memberikan nilai edukasi bagi pengunjung.
- Topik kelima menggambarkan Benteng Rotterdam sebagai destinasi wisata sejarah yang strategis karena berada di pusat kota dan memiliki kondisi bangunan yang masih baik. Selain itu, adanya kata "gratis" menunjukkan bahwa biaya kunjungan yang terjangkau turut menjadi pertimbangan wisatawan.
- Topik keenam secara khusus membahas biaya masuk wisata dan museum La Galigo. Wisatawan menilai biaya masuk relatif murah, meskipun terdapat biaya tambahan untuk memasuki area museum tertentu. Temuan ini menunjukkan bahwa faktor harga tetap menjadi perhatian dalam pengalaman wisata pengunjung.



Gambar 9. Visualisasi *Intertopic Distance Map*

Berdasarkan beberapa *insight* yang diperoleh maka dapat diketahui bahwa ulasan wisatawan terhadap Benteng Rotterdam tidak hanya berfokus pada aspek sejarah, tetapi juga mencakup kondisi lingkungan, lokasi strategis, fasilitas sekitar, pengalaman sosial, serta biaya kunjungan. Selain itu, visualisasi *Intertopic Distance Map* pada Gambar 9 menunjukkan terdapat sedikit irisan antara lingkaran Topik 1 dan Topik 2 yang mengindikasikan adanya keterkaitan makna antar topik. Kedekatan ini mengindikasikan bahwa kedua topik memiliki pola kemunculan yang relatif serupa dalam korpus ulasan dan sering muncul pada dokumen yang sama. Secara substantif, Topik 1 berkaitan dengan lokasi dan fasilitas di sekitar Benteng Rotterdam, sedangkan Topik 2 berkaitan dengan interaksi sosial melalui aktivitas belajar bahasa Inggris. Kedua topik tersebut sama-sama merepresentasikan pengalaman wisata pendukung yang diperoleh pengunjung di kawasan sekitar Benteng Rotterdam, sehingga memiliki kedekatan yang lebih tinggi dibandingkan topik lainnya Sebaliknya, Topik 4, Topik 5, dan Topik 6 berada pada posisi yang relatif berjauhan, menunjukkan bahwa topik-topik tersebut memiliki karakteristik pembahasan yang lebih spesifik dan berbeda dari dua Topik 1 dan 2. Selain itu, ukuran lingkaran

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

Topik 1 merupakan yang terbesar dibandingkan topik lainnya, yang menunjukkan bahwa topik tersebut memiliki proporsi kemunculan tertinggi dalam korpus ulasan. Hal ini menunjukkan bahwa lokasi strategis Benteng Rotterdam dan keberadaan berbagai fasilitas pendukung di sekitarnya merupakan aspek yang paling banyak dibahas oleh wisatawan dalam ulasan daring.

3.6. Pembahasan

Hasil pemodelan topik menggunakan metode Latent Dirichlet Allocation (LDA) menunjukkan bahwa Benteng Rotterdam memiliki beragam daya tarik yang tercermin dari ulasan wisatawan pada TripAdvisor Indonesia. Daya tarik yang paling dominan berkaitan dengan aspek sejarah dan kondisi fisik bangunan. Kata-kata seperti museum, Belanda, Diponegoro, sejarah, bangunan, bersih, dan indah menunjukkan bahwa wisatawan memandang Benteng Rotterdam sebagai destinasi wisata sejarah yang memiliki nilai edukatif sekaligus estetika visual. Selain itu, kondisi bangunan yang dinilai terawat dan menarik untuk aktivitas fotografi memperkuat citra objek wisata ini sebagai objek wisata budaya yang nyaman dikunjungi.

Selain aspek sejarah, hasil *topic modeling* juga menunjukkan bahwa lokasi strategis Benteng Rotterdam menjadi salah satu daya tarik penting bagi wisatawan. Kemunculan kata-kata seperti Pantai Losari, hotel, pusat kota, dan toko mengindikasikan bahwa wisatawan mempertimbangkan kemudahan akses serta keberadaan fasilitas pendukung di sekitar kawasan wisata. Hal ini menunjukkan bahwa keterhubungan dengan destinasi lain dan fasilitas penunjang memberikan pengalaman wisata yang lebih memuaskan bagi pengunjung. Temuan ini memperlihatkan bahwa daya tarik suatu destinasi tidak hanya ditentukan oleh objek utama, tetapi juga oleh lingkungan dan infrastruktur pendukung di sekitarnya.

Hasil penelitian juga mengungkap adanya daya tarik sosial dan ekonomi yang muncul dalam ulasan wisatawan. Interaksi dengan wisatawan asing menjadikan kawasan Benteng Rotterdam dimanfaatkan sebagai tempat belajar dan berlatih bahasa Inggris, sehingga objek wisata ini tidak hanya berfungsi sebagai tempat rekreasi tetapi juga sebagai ruang interaksi budaya. Di sisi lain, topik mengenai biaya masuk menunjukkan bahwa harga tiket yang relatif murah memberikan kesan positif terhadap pengalaman wisatawan. Faktor biaya yang terjangkau menjadi nilai tambah yang dapat meningkatkan minat kunjungan, terutama bagi wisatawan lokal dan pelajar.

Secara keseluruhan, hasil penelitian menunjukkan bahwa daya tarik utama Benteng Rotterdam terdiri atas nilai sejarah, suasana dan kondisi bangunan, lokasi strategis, interaksi sosial, serta biaya kunjungan yang terjangkau. Hal ini membuktikan bahwa metode LDA mampu mengidentifikasi persepsi dan pengalaman wisatawan secara efektif melalui analisis ulasan daring. Informasi yang diperoleh dapat dimanfaatkan oleh pengelola wisata untuk menyusun strategi promosi yang lebih tepat sasaran dengan menonjolkan aspek-aspek yang paling banyak diapresiasi oleh pengunjung.

Berdasarkan aspek kualitas pemodelan, visualisasi *Intertopic Distance Map* menunjukkan bahwa sebagian besar topik memiliki tema yang berbeda, meskipun terdapat kedekatan semantik antara Topik 3, Topik 4, dan Topik 5 yang sama-sama berkaitan dengan aspek sejarah Benteng Rotterdam. Kemunculan kata-kata seperti museum dan bangun pada beberapa topik menunjukkan adanya tema yang saling beririsan. Namun demikian, Topik 3 lebih menekankan kondisi fisik bangunan, Topik 4 berfokus pada sejarah kolonial dan Pangeran Diponegoro, sedangkan Topik 5 menggambarkan Benteng Rotterdam sebagai bangunan bersejarah. Kemiripan tersebut menunjukkan bahwa beberapa aspek daya tarik wisata Benteng Rotterdam masih memiliki keterkaitan yang kuat sehingga belum sepenuhnya terpisah secara semantik oleh model LDA. Hal ini juga tercermin dari nilai *coherence score* tertinggi 0,397836 yang mengindikasikan bahwa kualitas pemisahan topik masih dapat ditingkatkan.

Journal Software, Hardware and Information Technology SHIFT
ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

Meskipun penelitian ini berhasil mengidentifikasi daya tarik utama Benteng Rotterdam melalui pemodelan topik menggunakan metode LDA, penelitian ini masih memiliki beberapa keterbatasan. *Dataset* yang digunakan hanya berasal dari satu objek wisata, terbatas pada ulasan berbahasa Indonesia, dan hanya mengambil ulasan dari satu sumber, sehingga hasil penelitian belum dapat digeneralisasikan untuk destinasi wisata lain maupun ulasan multibahasa. Selain itu, metode LDA yang berbasis *bag-of-words* belum mampu memahami konteks semantik dan hubungan antar kata secara mendalam, sehingga beberapa topik yang dihasilkan masih memiliki keterkaitan makna yang berdekatan. Proses interpretasi dan penamaan topik juga masih dilakukan secara manual berdasarkan kata dominan dan visualisasi *word cloud*, sehingga berpotensi menimbulkan subjektivitas peneliti. Penelitian ini juga hanya berfokus menemukan topik pada ulasan pengguna, sehingga persepsi positif maupun negatif wisatawan pada setiap topik belum dapat diidentifikasi secara spesifik.

4. KESIMPULAN

Penelitian ini berhasil menerapkan *topic modeling* dengan metode Latent Dirichlet Allocation (LDA) untuk menemukan topik-topik utama pada ulasan wisatawan Benteng Rotterdam. Berdasarkan hasil pemodelan dan interpretasi topik melalui kata-kata dominan serta visualisasi topik, diperoleh enam daya tarik utama Benteng Rotterdam, yaitu nilai sejarah bangunan, suasana dan kondisi bangunan, lokasi strategis, pengalaman sosial dan budaya, fasilitas pendukung, serta biaya kunjungan yang terjangkau. Hasil penelitian menunjukkan bahwa metode LDA dapat digunakan untuk mengidentifikasi persepsi wisatawan terhadap objek wisata berdasarkan ulasan daring. Temuan penelitian ini dapat dimanfaatkan oleh pengelola wisata sebagai sumber informasi dalam menentukan daya tarik utama yang layak dipromosikan untuk meningkatkan minat kunjungan wisatawan, sekaligus sebagai dasar dalam pengembangan strategi promosi dan peningkatan kualitas layanan wisata berbasis ulasan pengunjung.

Penelitian ini masih memiliki keterbatasan pada penggunaan dataset yang hanya berasal dari satu objek wisata dan satu sumber ulasan, serta penggunaan metode LDA yang belum mampu memahami konteks semantik secara mendalam. Oleh karena itu, penelitian selanjutnya dapat menggunakan dataset yang lebih beragam, mengkombinasikan *topic modeling* dengan analisis sentimen, serta menerapkan metode berbasis semantik untuk memperoleh pemahaman yang lebih komprehensif mengenai pengalaman wisatawan.

DAFTAR PUSTAKA

[1] S. Ihsannudin and L. Dewi, "Efektivitas Media Sosial Youtube Sebagai Sumber Informasi Wisata Kuliner Jakarta Bagi Wisatawan," *Media Bina Ilm.*, vol. 17, no. 6, pp. 1133–1140, 2023.

[2] K. K. Kurniasari, E. M. Widanti, and A. N. Hidayah, "Literatur review: Pengaruh media sosial terhadap industri pariwisata dari perspektif destinasi dan wisatawan," *Indones. J. Tour. Bus. Entrep.*, vol. 1, no. 1, pp. 71–82, 2024.

[3] M. Rahman and M. N. Mia, "The influence of electronic word-of-mouth (eWOM) on promoting sustainable tourism in Bangladesh," *Hum. Behav. Emerg. Technol.*, vol. 2025, no. 1, p. 6650724, 2025.

[4] J. P. Atmaja, "Peran Teknologi Informasi Dalam Peningkatan Daya Saing Destinasi Pariwisata Di Indonesia," *J. Destin. Pariwisata*, vol. 11, no. 1, p. 151, 2023.

[5] S. D. Hawa and A. Naimah, "Pengaruh Destination Image, E-wom, Kualitas Pelayanan, Digital Marketing, dan Fasilitas Terhadap Keputusan Berkunjung Wisata Alam Sirah Pemali," *J. Bisnis, Manajemen, dan Akunt.*, vol. 12, no. 2, pp. 60–91, 2025.

[6] S. H. Dethan, B. D. Widiyasti, P. A. Wardi, and R. Ramdani, "Peran Digital Marketing Dan E-Wom (Electronic Word Of Mouth) Terhadap Keputusan Berkunjung Objek Wisata

ISSN 2779-8961 (online)
2808-3385 (print)
Vol. 6 No. 2 Tahun 2026
doi: https://doi.org/10.24252/shift.v6i2.298

Kabupaten Lombok Barat," *Econ. J. Ekon. Dan Bisnis*, vol. 2, no. 1, pp. 80–88, 2025.
[7] G. Ilieva and D. Boteva, "Effects of electronic word-of-mouth (eWOM) drivers on customers' behaviour.," *East. J. Eur. Stud.*, vol. 14, 2023.
[8] N. Cahyono, *Fundamental Topic Modelling: Memahami Tren & Topik dalam Text Mining*. MEGA PRESS NUSANTARA, 2024.
[9] M. Karabacak and K. Margetis, "Natural language processing reveals research trends and topics in The Spine Journal over two decades: a topic modeling study," *Spine J.*, vol. 24, no. 3, pp. 397–405, 2024.
[10] L. K. Bayu, "TA: Analisis Topic Modelling Ibukota Nusantara pada Platform X Menggunakan Latent Dirichlet Allocation," 2025, *Universitas Dinamika*.
[11] K. Watanabe and A. Baturo, "Seeded sequential LDA: A semi-supervised algorithm for topic-specific analysis of sentences," *Soc. Sci. Comput. Rev.*, vol. 42, no. 1, pp. 224–248, 2024.
[12] A. Abdelrazek, Y. Eid, E. Gawish, W. Medhat, and A. Hassan, "Topic modeling algorithms and applications: A survey," *Inf. Syst.*, vol. 112, p. 102131, 2023.
[13] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," *arXiv Prepr. arXiv2203.05794*, Mar. 2022, [Online]. Available: http://arxiv.org/abs/2203.05794
[14] I. Z. Mustaqim and R. R. Suryono, "A Systematic Literature Review of Topic Modeling Techniques in User Reviews.," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 11, no. 2, 2025.
[15] G. N. Buhari, D. Pramitasari, and A. Saifullah, "Persepsi Wisatawan terhadap Kualitas Produk Wisata: Fort Rotterdam, Di Makassar," *Sinektika J. Arsit.*, vol. 19, no. 1, pp. 96–106, 2022.
[16] M. A. Rajab, "Daya tarik wisata terhadap kepuasan wisatawan objek wisata Benteng Rotterdam Kota Makassar," *Pringgitan*, vol. 1, no. 2, 2020.
[17] A. Majid and S. Rabiah, "Communication Governance Strategy for the Cultural Conservation Preservation Center of South Sulawesi Province in Increasing Tourist Visits Fort Rotterdam Makassar," *RESPON J. Ilm. Mhs. ILMU Komun.*, vol. 3, no. 2, pp. 63–79, 2022.
[18] R. C. Sanco'o and M. Bisyri, "Strategi Komunikasi BPCB Sulsel Dalam Menarik Minat Kunjungan Wisatawan Mancanegara Terhadap Benteng Fort Rotterdam," *J. Komun.*, vol. 15, no. 1, pp. 26–35, 2022.
[19] R. Rehurek, "Gensim," 2019, *Python Software Foundation*: 4.4.0. [Online]. Available: https://pypi.org/project/gensim/4.4.0/
[20] Python Software Foundation, "re," 2001, *The Python Software Foundation*: 3.14. [Online]. Available: https://docs.python.org/3/library/re.html
[21] N. Project, "NLTK," 2025, *NLTK Project*: 3.9.2. [Online]. Available: https://www.nltk.org/
[22] A. Librian, "Sastrawi," 2015, v1.2.0. [Online]. Available: https://github.com/sastrawi/sastrawi
[23] S. J. Weston, I. Shryock, R. Light, and P. A. Fisher, "Selecting the number and labels of topics in topic modeling: A tutorial," *Adv. Methods Pract. Psychol. Sci.*, vol. 6, no. 2, p. 25152459231160104, 2023.
[24] L. Chen, J. Yan, J. Chen, Y. Sheng, Z. Xu, and M. Mahmud, "An event based topic learning pipeline for neuroimaging literature mining," *Brain Informatics*, vol. 7, no. 1, p. 18, 2020.
[25] J. Lee, J.-H. Kang, S. Jun, H. Lim, D. Jang, and S. Park, "Ensemble modeling for sustainable technology transfer," *Sustainability*, vol. 10, no. 7, p. 2278, 2018.