

PEMODELAN TOPIK ISU PARWISATA PULAU LOMBOK MENGGUNAKAN FOCUS WEB CRAWLER DAN LATENT DIRICHLET ALLOCATION

Supardianto¹, Seriasih², Lalu Mutawalli³

^{1,2}Informatika, Universitas Teknologi Mataram, ³Sistem Informasi, STMIK Lombok

Jln. Kampus Universitas Teknologi Mataram, Kekalik Mataram, 621939232

Jln. Basuki Rahmat No.105 Praya Lombok Tengah 83511

¹odyputra88@gmail.com, ²sriasish84@gmail.com, ³laluallistilo@gmail.com

Abstract

This research explores the potential of tourism on Lombok Island as a key driver of local economic, social, and cultural development. As a popular tourist destination, Lombok is known for its natural beauty which includes beaches and mountains as well as the rich culture of the Sasak tribe. This research uses topic modeling methods such as TF-IDF and Latent Dirichlet Allocation (LDA) to analyze social media content and news articles, which identified four to six main themes related to Lombok tourism, including destination promotion, tourist village development, and the impact of sporting events, such as MotoGP. The analysis results show that the LDA model has strong predictive ability with a perplexity value of -9.345 and a Coherence Score of 0.5873, reflecting good thematic coherence. The model for social media comments showed a perplexity of -6.276 and a Coherence Score of 0.513, while the model for online news articles resulted in a perplexity of -8.52. Overall, this research highlights the importance of tourism not only as a source of income, but also as a reinforcement of the social and cultural life of local communities, and provides relevant insights for the development of tourism that is responsive to the needs of tourists.

Keywords : Lombok Tourism, Topic Modeling, TF-IDF, Latent Dirichlet Allocation;

Abstrak

Penelitian ini mengeksplorasi potensi pariwisata di Pulau Lombok sebagai pendorong utama perkembangan ekonomi, sosial, dan budaya lokal. Sebagai destinasi wisata yang populer, Lombok dikenal dengan keindahan alamnya yang meliputi pantai dan gunung serta kekayaan budaya suku Sasak. Penelitian ini menggunakan metode pemodelan topik seperti TF-IDF dan Latent Dirichlet Allocation (LDA) untuk menganalisis konten media sosial dan artikel berita, yang mengidentifikasi empat hingga enam tema utama terkait pariwisata Lombok, termasuk promosi destinasi, pengembangan desa wisata, dan dampak acara olahraga, seperti MotoGP. Hasil analisis menunjukkan bahwa model LDA memiliki kemampuan prediksi yang kuat dengan nilai perplexity -9.345 dan Coherence Score 0.5873, mencerminkan koherensi tematik yang baik. Model untuk komentar media sosial menunjukkan perplexity -6.276 dan Coherence Score 0.513, sedangkan model pada artikel berita online menghasilkan perplexity -8.52. Secara keseluruhan, penelitian ini menyoroti pentingnya pariwisata tidak hanya sebagai sumber pendapatan, tetapi juga sebagai penguat kehidupan sosial dan budaya masyarakat lokal, serta memberikan wawasan yang relevan bagi pengembangan pariwisata yang responsif terhadap kebutuhan wisatawan.

Kata kunci : Pariwisata Lombok, Pemodelan Topik, TF-IDF, Latent Dirichlet Allocation

1. PENDAHULUAN

Pulau Lombok merupakan destinasi wisata populer di Indonesia, menawarkan wisata pantai, gunung, dan kekayaan budaya suku Sasak [1]. Pariwisata Lombok memberikan dampak

signifikan terhadap pertumbuhan transportasi darat dan udara [2]. Perkembangan infrastruktur pariwisata mempermudah aksesibilitas bagi wisatawan serta memberikan manfaat bagi penduduk lokal, terutama dalam bidang

ekonomi, sosial, dan budaya [3]. Pariwisata menciptakan lapangan pekerjaan di sektor transportasi, akomodasi, dan kerajinan lokal, serta mendorong pertumbuhan sektor perdagangan dan industri lokal. Secara sosial, pariwisata mempererat interaksi antara wisatawan dan penduduk lokal, memperkaya kehidupan sosial dan budaya Lombok.

Pemahaman mendalam tentang aspek destinasi, fasilitas, event dan layanan terkait pariwisata sangat penting untuk pengelolaan yang efektif dan berkelanjutan. Destinasi wisata menjadi daya tarik utama, sementara fasilitas, event social dan budaya mendukung wisatawan menjadi terhibur dan nyaman. Analisis isu-isu pariwisata terkini, terutama dari perspektif wisatawan, diperlukan untuk memberikan wawasan bagi pengembangan pariwisata yang responsif terhadap kebutuhan wisatawan.

Untuk memahami topik pariwisata Lombok, pendekatan yang digunakan meliputi pengumpulan data dari media sosial dan situs berita seperti Situs berita online lokal dan nasional. teknik *Focus Web Scraping* digunakan untuk mengumpulkan konten, postingan, artikel dan engegment, di Analisa dan dimodelkan menggunakan *Term Frequency Inverse Document Frequency (TF-IDF)* dan *Latent Dirichlet Allocation (LDA)*.

Beragam pendekatan telah dilakukan untuk analisis dan pemodelan isu pariwisata. Penelitian oleh [4] berfokus pada sistem yang memprediksi rating, mengumpulkan umpan balik, dan merekomendasikan lokasi wisata sesuai preferensi pengguna, menggunakan data dari ulasan, blog, dan rating di *Tripadvisor* dan *Yelp*. Metode *Random Forest*, *Decision Tree*, dan *TF-IDF* digunakan, menghasilkan model klasifikasi akurat serta rekomendasi topik yang baik.

Sementara itu, penelitian oleh [5] mengisi kesenjangan dalam studi space syntax dengan menerapkan LDA untuk mengidentifikasi topik terkait aksesibilitas pejalan kaki, yang meningkatkan pengalaman wisatawan di destinasi.

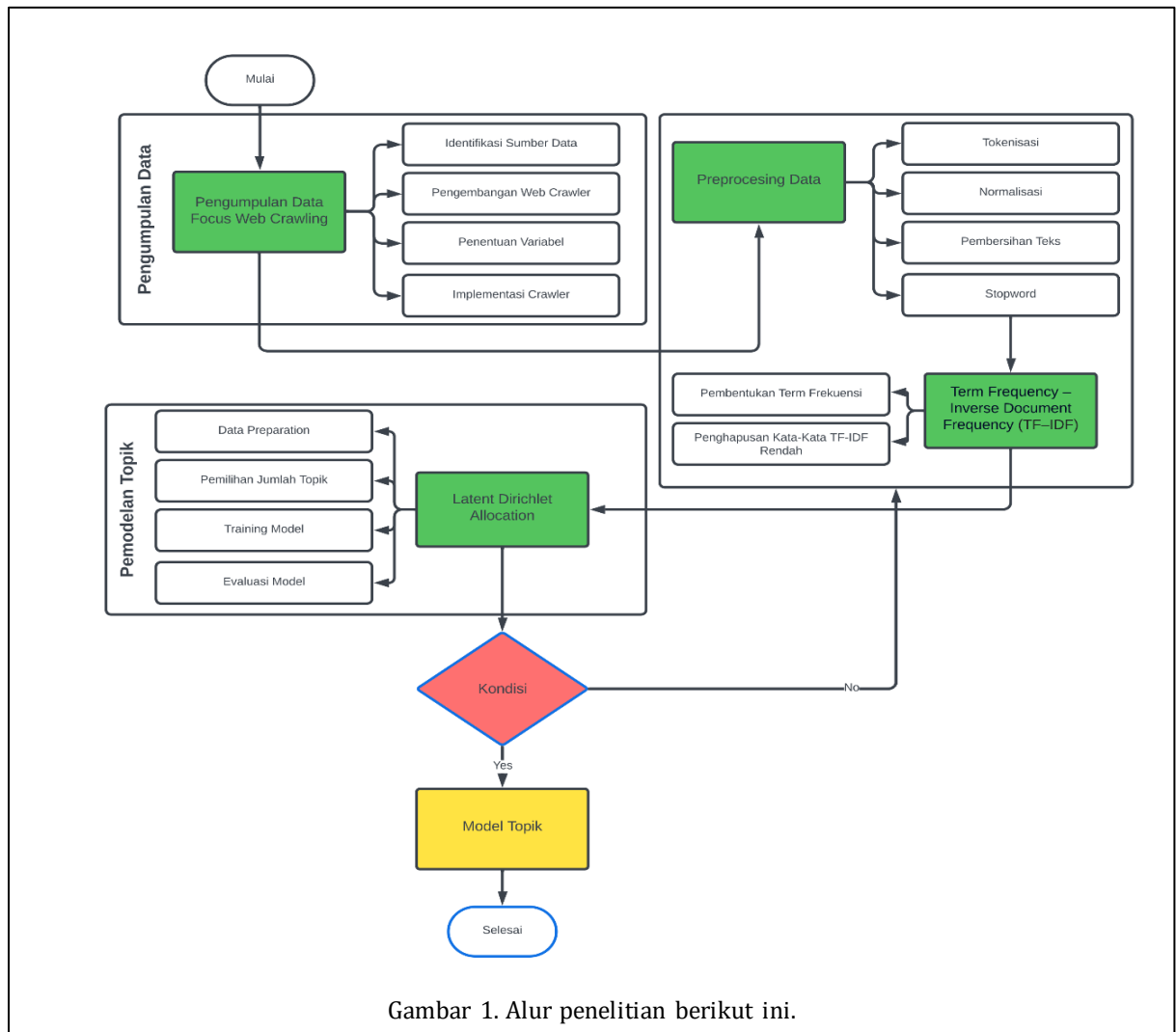
Penelitian oleh [6] menggunakan data ulasan dari *TripAdvisor* dan kedatangan wisatawan ke Hongkong dengan LDA untuk ekstraksi fitur, dan mengklasifikasikannya menggunakan *ANN*, *SVM*, dan *Random Forest*. Hasilnya menunjukkan integrasi topik dari ulasan media sosial dapat meningkatkan akurasi prediksi kunjungan wisatawan. Selanjutnya, penelitian [7] meneliti pengaruh *key opinion leader (KOL)* terhadap keputusan perjalanan wisatawan, dengan LDA yang menunjukkan bahwa informasi dari KOL menjadi referensi penting bagi wisatawan.

Penelitian oleh [8] mengklasifikasi opini wisatawan tentang pariwisata Lombok dengan menggunakan data dari *Twitter*, menerapkan *LDA* dan *Naive Bayes*, dan mencapai skor klasifikasi 0.92 dengan koherensi tertinggi 0.613 untuk kelas positif dan 0.528 untuk kelas negatif. Penelitian selanjutnya akan menerapkan metode web crawling di media sosial seperti Instagram dan Facebook, serta portal berita online seperti Detik dan Lombok Post, lalu menganalisis data tersebut menggunakan *TF-IDF* dan *LDA* untuk pemodelan topik. Hasil analisis ini memberikan wawasan tentang preferensi wisatawan (*netizen*), dan pengelola wisata di Lombok dapat menjadi panduan penmangku kebijakan pariwisata Lombok sesuai harapan wisatawan.

2. METODOLOGI PENELITIAN

2.1. Skema Alur Penelitian

Penelitian ini bertujuan untuk menganalisis model pengetahuan model topik isu pariwisata Lombok. Alur penelitian terdiri dari empat tahap, yaitu pengumpulan data dengan focus web scraping, preprocessing teks, penghapusan stopword dengan TF-IDF, dan pemodelan topik dengan LDA. Setiap tahap akan dijelaskan secara lebih rinci dalam gambar 2. Alur penelitian berikut ini.



2.2. Pengumpulan Data

Dalam penelitian ini, pengumpulan data dilakukan dengan Teknik scrapy bersumber dari platform media sosial Facebook dan Instagram, dengan fokus pada akun-akun yang terkait dengan pariwisata di wilayah Lombok. Selain data yang berasal dari facebook dan Instagram penelitian ini juga memanfaatkan data yang diperoleh dari beragam sumber informasi, khususnya portal berita online yang relevan dengan konteks yang diteliti.

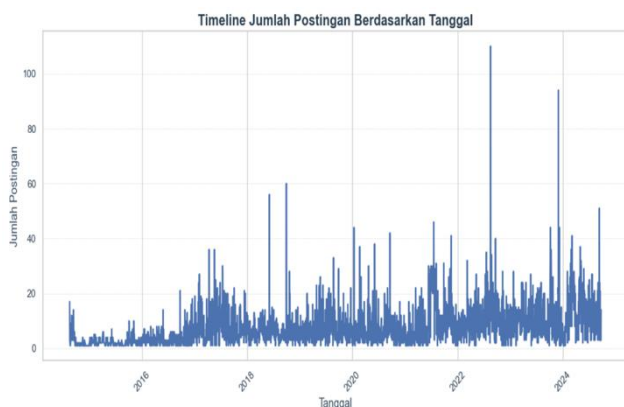
TABEL II. SUMBER DATA SOSIAL MEDIA INSTAGRAM, INSTAGRAM, DAN WEB BERITA ONLINE

	Akun Instagram	Akun Facebook
Spsial Media	@disparlomboktimur, @gomandalika, @pariwisata.lombokbarat, @dispar_mataram, @lombokutara.tourism, @wonderfullomboksumbawa, @poltekiparlombok, @lomboksumbawa.go, @InfoRinjani, @pearloftrawangan, dan @thegiliway.	WonderfulLombokSumbawa, gomandalika, disparlobar, lombokutara.tourism, LombokTraveler official, rinjanilombokgeopark, Gili.Trawangan. Lombok.Indonesia, gilitrawanganlombokok, dan

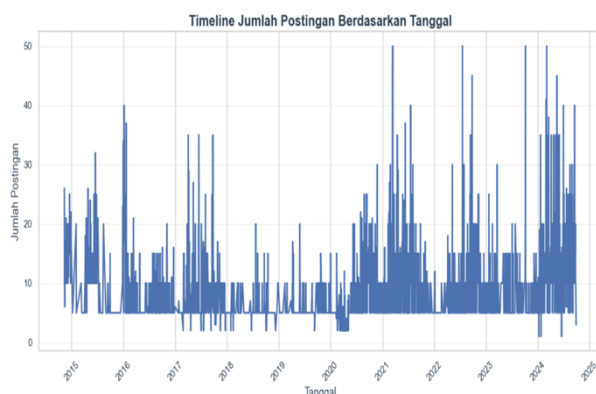
	Lolita-Gili-Trawangan.
Portal Berita Online	lombokpost.jawapos.com/wisata, poltekpar-lombok-today, lombok.tribunnews, suara NTB, dan detik.

Akun-akun pada table 1. memberikan perspektif yang beragam tentang aktivitas dan promosi pariwisata di Lombok, serta interaksi masyarakat dengan destinasi-destinasi wisata yang ditawarkan. Selain itu, penelitian ini juga memanfaatkan data dari beberapa portal berita online terkemuka yang membahas pariwisata, termasuk Penggunaan sumber data yang beragam ini bertujuan untuk menghasilkan analisis yang komprehensif mengenai dinamika pariwisata di Lombok, serta untuk memahami persepsi dan perilaku wisatawan.

Data yang dikumpulkan yang bersumber dari instagram sejak tahun 2016 hingga 2024 sebanyak 26457 postingan dan komentar, dapat dipantau melalui Gambar 1. Time line jumlah postingan di sosial media Instagram, dan Gambar 2. Berasal dari facebook yang di visualisasi berdasarkan tanggal.

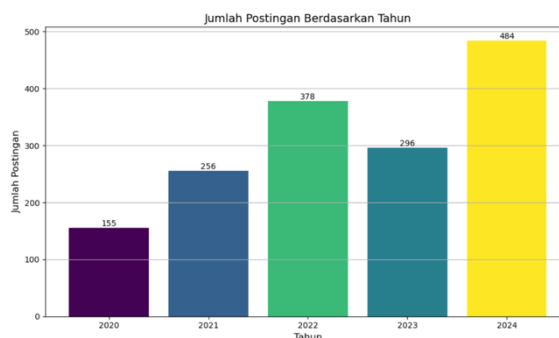


Gambar 1. Timeline Data Instagram



Gambar 2. Timeline Data Instagram

Penggunaan berbagai sumber ini dimaksudkan untuk memperkaya perspektif dan memberikan konteks yang lebih luas dalam analisis yang dilakukan. Jumlah total artikel yang telah didapatkan pada semua portal berita online yang menjadi objek penelitian ini mencapai 1.569 artikel. Data ini mencerminkan tingkat keterlibatan dan aktivitas media dalam meliputi isu-isu yang berkaitan dengan pariwisata dan kebudayaan di Lombok. Pada Gambar 3. Menampilkan timeline data yang terkumpul, menunjukkan distribusi jumlah artikel yang diterbitkan setiap tahunnya.



Gambar 3. Jumlah artikel yang telah dipublikasikan oleh Portal berita online terkait berita Pariwisata Lombok

2.3. Preprocessing Data

Mencakupi tokenisasi dengan memisahkan kata-kata atau frasa, normalisasi membuat huruf menjadi kecil semua, pembersihan teks hapus non-teks seperti URL, karakter khusus, dan tautan serta penghapusan stopwords identifikasi kata-kata umum yang tidak memberikan nilai tambah dalam analisis seperti “dan”, “atau”, dan kata stopwords.

2.4. Analisa Data TF-IDF

Tahapan ini bertujuan melanjutkan analisis tahap kata-kata yang paling banyak terdapat pada seluruh dokumen, dengan metode *WordCloud* dan *TF-IDF*, Penerapan metode TF-IDF pada ekstraksi kata dalam postingan, artikel, dan komentar social media dan berita membantu menguraikan kompleksitas teks dan mendapatkan informasi. TF mengukur frekuensi kemunculan kata dalam dokumen, yang penting dalam analisis data teks. (1).

$$tf(t, d) = \frac{f_{t,d}}{\sum t^l \in df_{f^l,d}} \quad (1)$$

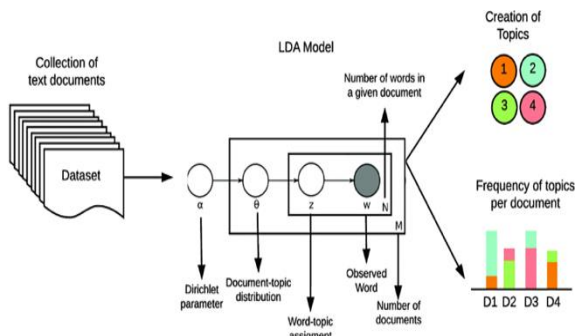
Sementara itu, IDF (2) bertujuan mengukur seberapa unik atau penting suatu kata dalam kumpulan dokumen[9]. Semakin jarang kata muncul di banyak dokumen, semakin tinggi nilainya, menunjukkan kata tersebut lebih bermakna.

$$idf(t, D) = \log \frac{N}{|\{d \in D: t \in d\}|} \quad (2)$$

2.5. Topik Modeling LDA

Dalam pemodelan topik untuk postingan konten, artikel, dan komentar netizen tentang pariwisata Lombok, kata-kata yang memiliki keterkaitan digabungkan menjadi topik-topik bermakna melalui metode LDA. Rumus LDA yang dijelaskan dalam (3) menjadi dasar matematis untuk membentuk topik berdasarkan pola kata-kata yang sering muncul bersama, sehingga memberikan wawasan mendalam tentang isu-isu yang dibahas oleh netizen mengenai pariwisata Lombok.

$$P(z_i = j | z_{-i}, w_i, d_i) \propto \frac{C_{w_{ij}}^{WT} + \eta}{\sum_{w=1}^W C_{wj}^{WT} + W\eta} \frac{C_{d_{ij}}^{DT} + \alpha}{\sum_{t=1}^T C_{dt}^{WT} + T\alpha} \quad (3)$$



Gambar. 4 Latent Dirichlet Allocation Model

Pada tahap LDA, teks dianalisis dengan memproses distribusi topik dalam dokumen menggunakan parameter Dirichlet, yang menunjukkan proporsi setiap topik. Kata-kata dalam dokumen diberikan ke topik tertentu melalui proses penugasan. Frekuensi kemunculan topik dihitung untuk membentuk topik berdasarkan distribusi[10].

2.6. Evaluasi Model

Pengujian pada model dilakukan untuk mengevaluasi kinerja model. Metode yang digunakan adalah *coherence score* dan *perplexity*. *Coherence* mengukur keterkaitan kata dalam satu topik [11] dan *perplexity* mengevaluasi seberapa baik model memprediksi data baru [12], [13]. Rumus coherence (3) dan perplexity (4).

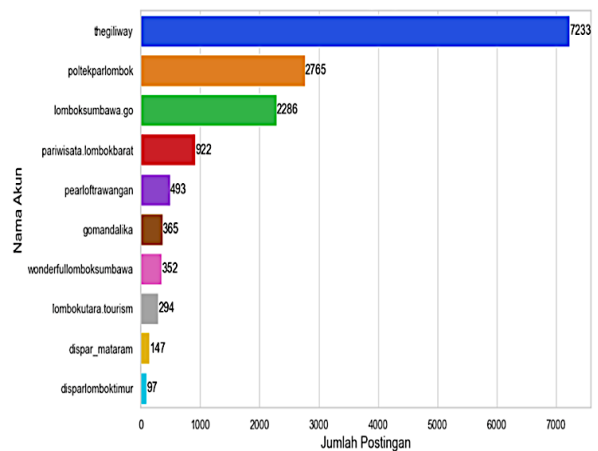
$$C_v = \frac{1}{M} \sum_{i=1}^{M-1} \sum_{j=i+1}^M \text{similarity}(w_i, w_j) \quad (3)$$

$$C_v = \exp \left(-\frac{1}{N} \sum_{d=1}^D \log P(w_d) \right) \quad (4)$$

3. HASIL DAN PEMBAHASAN

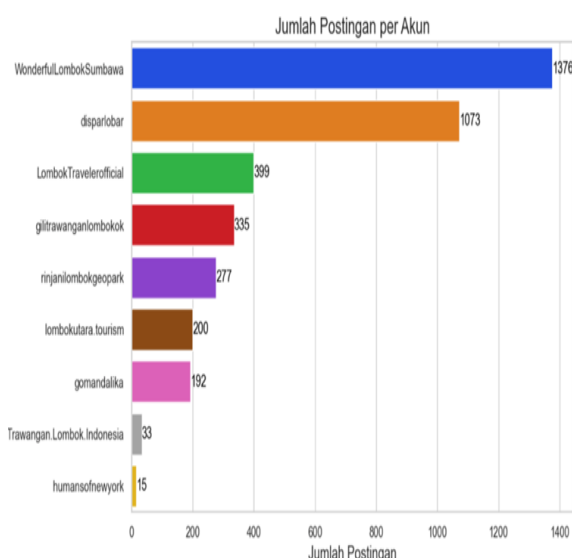
3.1 Eksplanatory Data

Setelah melakukan penghapusan duplikasi, data menjadi 18.045 postingan. Akun the giliway menjadi yang paling aktif dengan 7.319 postingan, diikuti oleh poltekpar lombok 3.444, lombok sumbawa 2.314 Akun-akun lain seperti pariwisata lombokbarat 992 dan pearlof trawangan 640 menunjukkan jumlah postingan yang lebih rendah. Akun dengan aktivitas paling sedikit adalah go mandalika 372, wonderful lombok sumbawa 363, lombok utara tourism 317, dispar lombok timur 185, dispar_mataram 153. Ini menunjukkan bahwa beberapa akun kurang aktif dalam berbagi konten secara konsisten di Instagram Gambar 5.



Gambar 5. Jumlah postingan pada setiap akun Instagram

Berdasarkan data yang diperoleh, akun facebook Wonderfull Lombok Sumbawa memiliki jumlah postingan tertinggi dengan 1.376 postingan, diikuti oleh dispar loyar dengan 1.073 postingan. Lombok Travel official berada di urutan ketiga dengan 399 postingan. Akun gilitrawangan dan rinjani lombok geopark masing-masing memiliki 335 dan 277 postingan. Partisipasi lebih rendah terlihat pada akun lombokutara tourism 200, go mandalika 192, dan Gili Trawangan Lombok Indonesia 33. Akun human sof newyork mencatat aktivitas postingan paling minimal, hanya 15 postingan. Gambar 6. Menunukan Analisa visual distribusi Jumlah Postingan Akun Facebook Terkait Pariwisata Lombok.

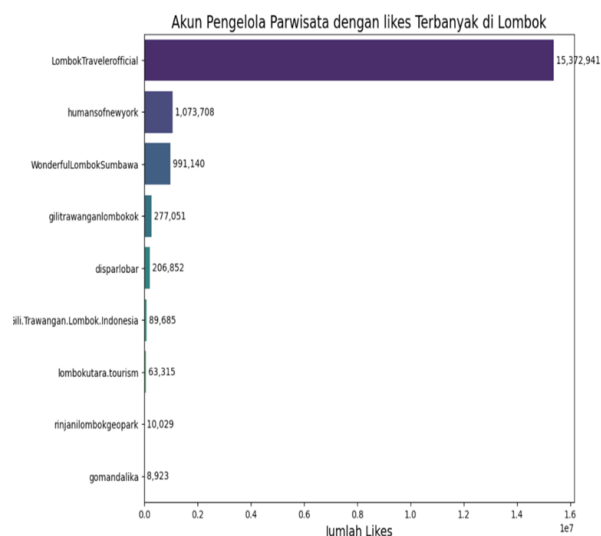


Gambar 6. Distribusi postingan pada akun facebook

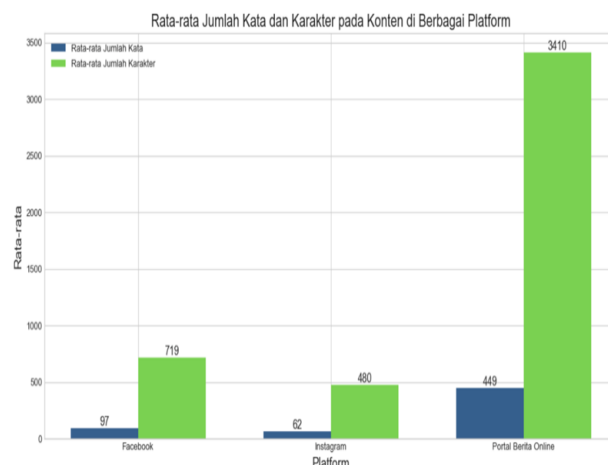
Berdasarkan data, akun Lombok Travel official menempati posisi tertinggi dengan 15.372.941 like, menunjukkan daya tarik besar kontennya dalam promosi pariwisata Lombok. Wonderfull Lombok Sumbawa berada di posisi kedua dengan 991.140 like, diikuti oleh humansofnewyork 1.073.708 like. Akun disparlobar dan gili trawangan lombok mencatat jumlah like signifikan, masing-masing 206.852 dan 227.052. Akun go mandalika, lombokutara tourism, dan rinjani lombok geopark menunjukkan pengaruh lebih terbatas dengan jumlah like lebih rendah.

Gambar 7. Akun facebook pengelola wisata yang paling mendapatkan like paling tinggi

Analisis menunjukkan perbedaan signifikan dalam jumlah kata dan karakter antara Facebook, Instagram, dan portal berita online. Facebook memiliki rata-rata 97 kata dan 719 karakter per posting, sementara Instagram lebih fokus pada



visual dengan 62 kata dan 480 karakter. Sebaliknya, portal berita online memiliki rata-rata 449 kata dan 3.410 karakter per artikel.

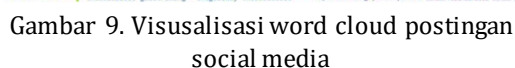


Gambar 8. Rata-Rata jumlah kata dan karakter pada konten di berbagai platform

3.2 Analisis Word Cloud Pada Konten Sosial Media, Artikel, dan Ulasan netizen

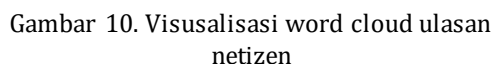
3.2.1 Konten Sosial Media

Analisis WordCloud menunjukkan bahwa kata "Lombok" muncul paling sering (39.424), menandakan fokus utama pada Lombok. Kata "Pariwisata" (19.283) dan "Wisata" (15.697) menyoroti sektor pariwisata, sementara "Desa" (13.302) dan "Masyarakat" (6.420) menunjukkan pentingnya peran lokal dalam pengembangan pariwisata. Geografi NTB juga menjadi fokus. Gambar 9. Visualisasi WordCloud Frekuensi Kata pada konten sosial media terkait pariwisata Lombok.



3.2.2 Ulasan Netizen

Analisis WordCloud ditunjukkan pada Gambar 10., pada komentar netizen di Instagram menunjukkan bahwa kata "gass" dan "nadiaa" mendominasi dengan frekuensi tertinggi 17.000. Kata-kata yang lebih relevan dengan pariwisata, seperti "lombok" 361, "gili" 304, dan "ntb" 176, mencerminkan minat terhadap destinasi wisata spesifik. Sentimen positif seperti "beautiful" 181, "love" 257, "amazing" 134, dan "great" 128 juga banyak muncul, menunjukkan persepsi wisatawan yang umumnya baik terhadap pariwisata di Lombok dan Nusa Tenggara Barat.



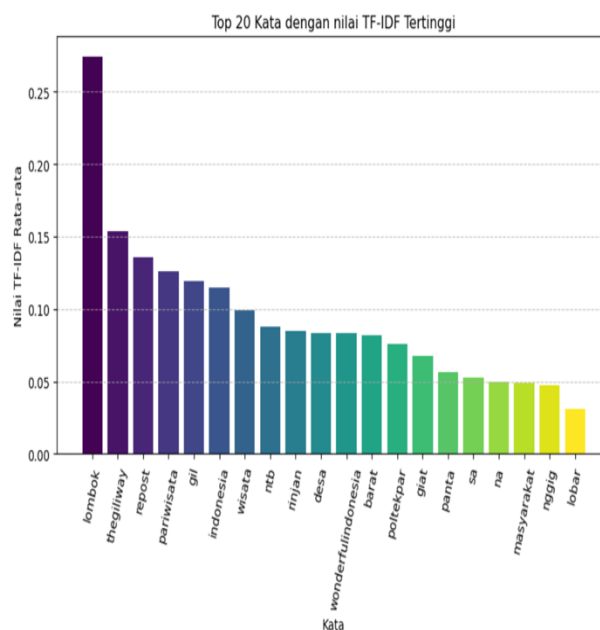
Analisis WordCloud menunjukkan frekuensi kata yang mendominasi konten terkait pariwisata Lombok. "Lombok" muncul 5.344 kali, menunjukkan fokus utama, sedangkan "Wisata" muncul 4.918 kali, menyoroti sektor pariwisata. Kata "Desa" 3.023 dan "Indonesia" 2.856 mengaitkan konten dengan desa dan identitas nasional. "Pariwisata" 2.292 dan "Mandalika" 2.278 mempertegas pentingnya Mandalika sebagai destinasi wisata.



3.3 Analisa TF-IDF

3.3.1 Analisis TF-IDF Konten Sosial Media

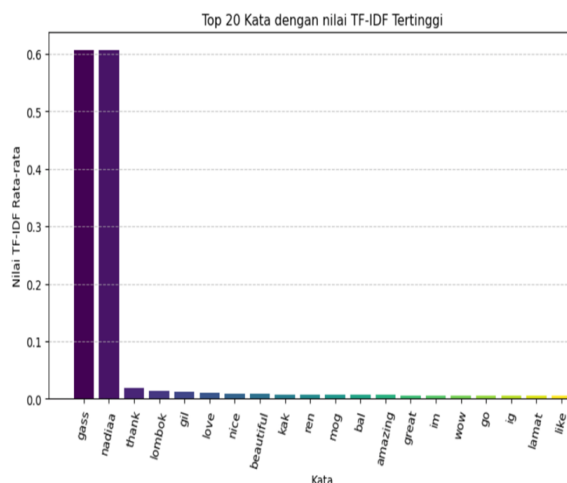
Analisis TF-IDF pada konten pariwisata Lombok menunjukkan bahwa kata "Lombok" memiliki bobot tertinggi 0.274, menandakan fokus wilayah tersebut. Kata "thegiliway" 0.153 mencerminkan perhatian terhadap Gili, sedangkan "repost" 0.136 dan "pariwisata" 0.126 menunjukkan relevansi konten. Istilah "Indonesia" 0.115 dan "wisata" 0.098 mengindikasikan konteks nasional, sementara "NTB" 0.088 dan "Rinjani" 0.085 menekankan promosi destinasi di Nusa Tenggara Barat. Gambar 12 merupakan hasil visualisasi analisis TF-IDF pada konten social media.



Gambar 12. Distribusi Kata- dengan Nilai TF-IDF dalam postingan Pariwisata Lombok

3.3.2 Analisis TF-IDF Ulsan Netizen

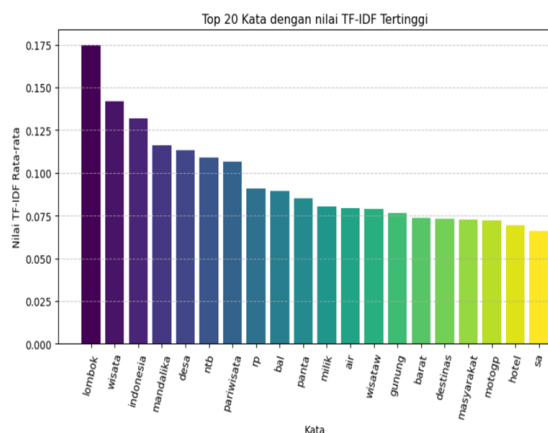
Analisis TF-IDF menunjukkan bahwa kata "gass" dan "nadia" memiliki nilai tertinggi (0.606), menandakan dominasi dalam percakapan netizen. Kata-kata lain seperti "thank," "lombok," "gil," "love," dan "beautiful" mencerminkan sentimen positif, sedangkan "great," "wow," dan "go" menunjukkan reaksi emosional serta interaksi melalui media sosial.



Gambar 13. Distribusi Kata- dengan Nilai TF-IDF dalam Ulasan Netizen isu Pariwisata Lombok

3.3.3 Analisis TF-IDF Web Berita Online

Analisis TF-IDF mengungkap relevansi kata dalam pemberitaan media online mengenai Lombok dan pariwisata. Kata "Lombok" memiliki nilai tertinggi (0.174940), diikuti "wisata" (0.141862) dan "indonesia" (0.132028). Istilah lain, seperti "mandalika" dan "desa," menyoroti aspek ekonomi dan sosial dalam narasi pariwisata.

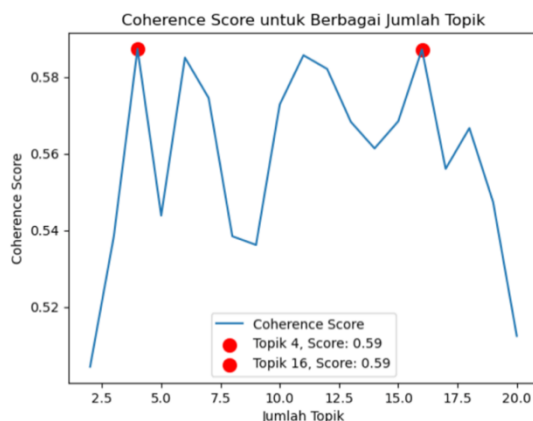


Gambar 14. Visualisasi Hasil analisis kata kunci pemberitaan pariwisata Lombok pada web portal berita online.

3.4 Topik Model LDA

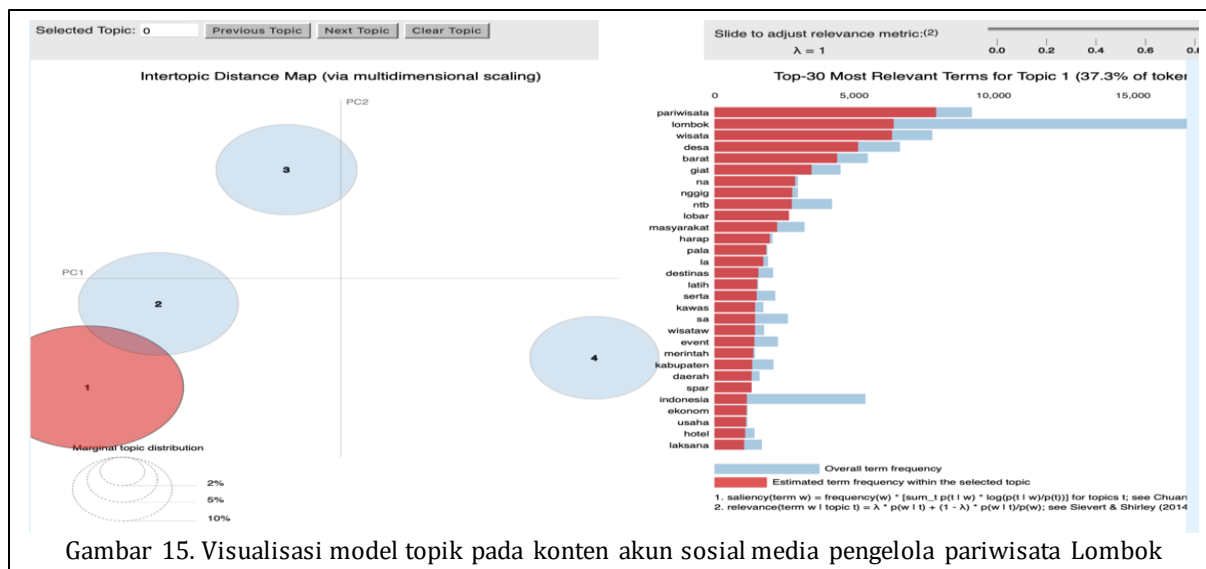
3.4.1 Topik model Konten Sosial Media

Dari hasil pemodelan topik, terlihat bahwa jumlah topik 4 (0.5873) dan 16 (0.5872) memiliki coherence score tertinggi, menunjukkan bahwa kedua model ini memberikan pemisahan topik yang lebih baik dibandingkan model lain[11]. Hasil pemodelan topik menunjukkan bahwa model dengan 4 (0.5873) dan 16 (0.5872) topik memiliki coherence score tertinggi, mengindikasikan pemisahan topik yang baik. Coherence score menilai sejauh mana topik dapat dipahami, dengan nilai lebih tinggi menunjukkan konsistensi. Model dengan 4 topik lebih sederhana, ideal untuk memahami topik dasar, sementara model 16 topik menawarkan detail lebih mendalam. Meski demikian, model 4 topik tetap menarik karena mudah dipahami tanpa mengorbankan kualitas interpretasi, sebagaimana ditunjukkan oleh coherence score yang tinggi.



Gambar 14. Perbandingan coherence score berdasarkan jumlah topik dalam model analisis konten pariwisata Lombok

Berdasarkan pemodelan topik pada konten akun pengelola pariwisata Lombok, terdapat empat topik utama yang teridentifikasi. Topik 0 didominasi oleh kata-kata seperti "lombok," "desa," "rinjani," "wisata," dan "pantai," yang menunjukkan fokus pada destinasi wisata alam, terutama desa-desa wisata, Gunung Rinjani, serta pantai-pantai di Lombok. Topik 1 berfokus pada pariwisata di Lombok Barat, dengan kata kunci seperti "pariwisata," "barat," "nggig," dan "lobar," mencerminkan kegiatan pariwisata lokal dan inisiatif masyarakat. Topik 2 mengedepankan pengembangan desa wisata dan pelatihan



pariwisata, di mana kata-kata seperti "desa," "pariwisata," "poltekpar," dan "latih" menunjukkan pentingnya pendidikan. Topik 3 menyoroti promosi wisata di Gili Trawangan dan kawasan sekitarnya, dengan kata-kata seperti "thegiliway," "repost," dan "gilitrawang," mencerminkan kolaborasi digital dalam promosi destinasi.

Tabel II. Topik model pada konten akun pengelola pariwisata lombok

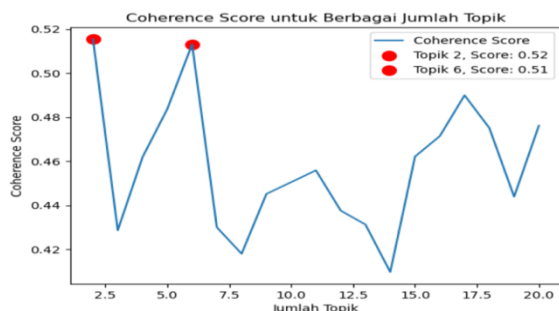
Topik	Konten Sosial Media
0	0.034*"lombok" + 0.009*"poltekpar" + 0.008*"panta" + 0.007*"wonderful indonesia" + 0.007*"desa" + 0.006*"air" + 0.006*"pulau" + 0.006*"sonaindonesia" + 0.006*"indonesia" + 0.006*"wisata"
1	0.017*"lombok" + 0.015*"rinjan" + 0.012*"indonesia" + 0.009*"geopark" + 0.006*"mbalun" + 0.006*"ntb" + 0.006*"lomba" + 0.006*"mandalika" + 0.006*"halal" + 0.005*"foto"
2	0.031*"thegiliway" + 0.021*"repost" + 0.019*"gilitrawang" + 0.019*"gil" + 0.019*"lombok" + 0.012*"giliisland" + 0.011*"thegalibible" + 0.008*"photo" + 0.008*"bal" + 0.007*"island"
3	0.026*"pariwisata" + 0.021*"lombok" + 0.021*"wisata" + 0.017*"desa" + 0.014*"barat" + 0.011*"giat" + 0.009*"na" + 0.009*"nggig" + 0.009*"ntb" + 0.009*"lobar"

Hasil pemodelan topik ini membagi kata-kata dominan menjadi empat topik utama terkait pariwisata Lombok. Topik 0 berfokus pada destinasi wisata alam dan budaya dengan kata

kunci seperti "pantai," "desa," dan "poltekpar." Topik 1 berkaitan dengan wisata pegunungan, dengan kata seperti "rinjani" dan "halal." Topik 2 mengangkat wisata Gili, menyoroti kata "thegiliway" dan peran media sosial dalam promosi. Topik 3 menekankan pariwisata terorganisir dan kegiatan lokal dengan kata "pariwisata" dan "lobar." Secara keseluruhan, model ini mencerminkan fokus pariwisata Lombok, dari alam hingga promosi lokal. Gambar 15 menunjukkan visualisasi model konten yang diunggah oleh pengelola pariwisata Lombok.

3.4.2 Topik Model Ulasan Netizen

Berdasarkan hasil pengujian dengan didasarkan pada keseimbangan antara coherence score dan interpretabilitas topik. Terdapat jumlah 2 topik memiliki coherence score tertinggi (0.515) namun topik terlalu sedikit sehingga tidak dapat merepresentasikan topik karena kurang menangkap keragaman data secara detail. Sebaliknya, topik dengan 6 memiliki coherence score sedikit lebih rendah (0.513), ini memberikan granularitas yang lebih baik, sehingga memungkinkan mengelompokan tema yang lebih sfesifik. Sedangkan dengan jumlah 7 hingga delapan dan seterusnya mengalami penurunan bahkan memperburuk kemampuan model dalam memsiah topik. Sehingga pemilihan jumlah topik sebanyak 6 dianggap paling optimal. Hasil visualisasi dari analisis coherence score ditunjukkan pada Gambar 16. Optimalisasi pemilihan jumlah topik terhadap komentar netizen terhadap Pariwisata Lombok.



Gambar 16. Optimalisasi pemilihan topik komentar netizen terhadap pariwisata Lombok

Hasil analisis model topik dari komentar netizen tentang pariwisata Lombok, pada table 2. Adalah topik utama hasil pemodelan komentar netizen merefleksikan topik-topik utama yang sering dibahas oleh pengguna sosial media, memberikan gambaran minat, tren, dan topik diskusi.

Tabel III. Topik model pada konten akun pengelola pariwisata lombok

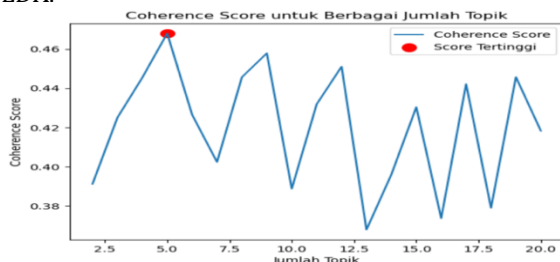
Topik	Konten Sosial Media
0	0.033*"gil" + 0.031*"bal" + 0.017*"thegiliway" + 0.016*"beach" + 0.014*"love" + 0.013*"like" + 0.012*"place" + 0.011*"back" + 0.010*"island" + 0.009*"u"
1	0.020*"sunset" + 0.018*"cute" + 0.018*"travel" + 0.017*"fun" + 0.016*"lfie" + 0.016*"fashion" + 0.016*"sun" + 0.016*"summer" + 0.016*"kis" + 0.016*"luxury"
2	0.024*"party" + 0.022*"bikin" + 0.014*"im" + 0.012*"need" + 0.011*"sharing" + 0.011*"get" + 0.010*"follow" + 0.008*"post" + 0.008*"hotel" + 0.006*"bal"
z	0.386*"gass" + 0.386*"nadiaa" + 0.008*"thank" + 0.006*"nice" + 0.005*"beautiful" + 0.004*"amazing" + 0.004*"love" + 0.004*"photo" + 0.003*"cool" + 0.003*"pic"
4	0.022*"lombok" + 0.016*"go" + 0.015*"kak" + 0.013*"ntb" + 0.012*"day" + 0.010*"pariwisat" + 0.010*"one" + 0.009*"mog" + 0.009*"let" + 0.008*"poltekpar"
5	0.018*"look" + 0.007*"t" + 0.007*"horse" + 0.006*"trawang" + 0.006*"yuk" + 0.005*"pretty" + 0.005*"orang" + 0.004*"te" + 0.004*"k" + 0.004*"cek"

Hasil intepretasi Tabel 2 adalah sebagai berikut. Topik 0 menggambarkan diskusi seputar lokasi wisata seperti "gil" (kemungkinan Gili Islands), "bal" (Bali), dan elemen destinasi pantai, seperti "beach", "island", dan pengalaman pengunjung yang menyukai destinasi ini. Topik 1 terkait dengan pengalaman wisata musim panas, di mana kata-kata seperti "sunset", "travel", "fun", dan "luxury" sering muncul, menunjukkan minat pada wisata yang berfokus pada keindahan alam dan liburan mewah. Topik 2 menunjukkan interaksi di media sosial yang berfokus pada aktivitas sosial, seperti "party", "bikin" (konten), serta promosi hotel dan layanan melalui kata "follow", "post", dan "hotel". Topik 3 didominasi oleh kata "gass" dan "nadiaa", yang menunjukkan adanya fenomena viral atau tren dalam media sosial terkait pengguna atau influencer tertentu. Topik 4 berfokus pada pariwisata Lombok dan NTB secara umum, dengan kata-kata seperti "lombok", "kak", "ntb", dan "pariwisat", mengindikasikan pembahasan mengenai destinasi lokal dan pendidikan pariwisata. Topik 5 melibatkan deskripsi visual yang lebih umum, dengan kata seperti "look", "pretty", "yuk", dan "horse", yang kemungkinan mencerminkan promosi visual dan ajakan untuk berkunjung. Hasil visualisasi topik model pada ulasan netizen dapat dilihat pada Lampiran 1, pada Gambar 17. Topik utama komentar netizen di Instagram terhadap Pariwisata Lombok

3.4.3 Topik Model Artikel pemberitaan online

Pemilihan jumlah topik yang paling optimal dari dataset artikel berita online diperlukan agar topik yang terbentuk memiliki landasan yang pasti. Pada penelitian ini diterapkan metode coherence score agar kualitas topik memiliki validasi yang baik. Hasil analisis kohence score dalam penelitian ini menunjukkan bahwa LDA dengan jumlah topik 5, hal demikian dikonfirmasi dengan nilai hasil koherensi tertinggi, yaitu 0.4679 rentang percobaan dari 2 hingga 20 topik. Nilai kohernsi menunjukkan peningkatan signifikan dari 2 hingga 5, yang mengindikasikan bahwa topik ini paling mampu menangkap makna semantic dari kata-kata dalam data artikel yang dipublikasikan oleh web berita online. Setelah mencapai puncak di 5 topik, terdapat penurunan yang jelas pada jumlah topik 10 dan 17, serta fluktuasi nilai pada jumlah topik berpotensi mengurango kejelasan makna. Gambar 10. Menyajikan hasil visualisasi analisis koherensi score yang digunakan untuk

menentukan jumlah topik optimal dalam model LDA.



Gambar 17. Visualisasi Koherensi Score Untuk Berbagai Jumlah Topik Artikel berita online

Hasil topik analisis topik menggunakan LDA dirangkum pada Tabel VI. Memberikan wawasan yang mendalam tentang tema yang muncul dalam artikel web berita online. Hasil pemodelan topik menunjukkan lima topik utama yang teridentifikasi dalam artikel web online, beserta kata-kata kunci dan bobotnya.

Tabel VI. Topik model pada konten akun pengelola pariwisata lombok

Topik	Konten Sosial Media
0	0.033*"gil" + 0.031*"bal" + 0.017*"thegiliway" + 0.016*"beach" + 0.014*"love" + 0.013*"like" + 0.012*"place" + 0.011*"back" + 0.010*"island" + 0.009*"u"
1	0.020*"sunset" + 0.018*"cute" + 0.018*"travel" + 0.017*"fun" + 0.016*"lfie" + 0.016*"fashion" + 0.016*"sun" + 0.016*"summer" + 0.016*"kis" + 0.016*"luxury"
2	0.024*"party" + 0.022*"bikin" + 0.014*"im" + 0.012*"need" + 0.011*"sharing" + 0.011*"get" + 0.010*"follow" + 0.008*"post" + 0.008*"hotel" + 0.006*"bal"
z	0.386*"gass" + 0.386*"nadiaa" + 0.008*"thank" + 0.006*"nice" + 0.005*"beautiful" + 0.004*"amazing" + 0.004*"love" + 0.004*"photo" + 0.003*"cool" + 0.003*"pic"
4	0.022*"lombok" + 0.016*"go" + 0.015*"kak" + 0.013*"ntb" + 0.012*"day" + 0.010*"pariwisat" + 0.010*"one" + 0.009*"mog" + 0.009*"let" + 0.008*"poltekpar"
5	0.018*"look" + 0.007*"t" + 0.007*"horse" + 0.006*"trawang" + 0.006*"yuk" + 0.005*"pretty" +

0.005*"orang" + 0.004*"te" + 0.004*"k" + 0.004*"cek"
--

Hasil analisis topik yang ditampilkan dalam Tabel VI menunjukkan beberapa fokus utama. Topik 0 berfokus pada pariwisata, dengan kata kunci seperti "desa," "Lombok," dan "Indonesia," menyoroti relevansi sektor pariwisata terhadap ekonomi dan sosial masyarakat. Ini menunjukkan pentingnya pariwisata sebagai pendorong utama manfaat bagi komunitas lokal. Topik 1 menekankan Mandalika dan MotoGP, mencerminkan minat terhadap acara olahraga internasional yang dapat meningkatkan profil wisata daerah. Topik 2 mengeksplorasi keindahan alam Lombok dengan kata kunci "gunung," "pantai," dan "Rinjani." Topik 3 menyoroti pengalaman relaksasi yang berkaitan dengan keindahan laut. Topik 4 mengaitkan Mandalika dengan sektor perhotelan dan infrastruktur akomodasi. Hasil pemodelan topik disajikan dalam visualisasi pada Gambar 18, yang menggambarkan distribusi topik dan keterkaitan istilah dominan. Topik 5, ditandai dengan warna merah, memiliki lingkaran terbesar, menunjukkan dominasi dalam distribusi, dengan istilah seperti "wisata," "NTB," dan "Lombok" mendominasi. Hasil visualisasi topik model pada ulasan netizen dapat dilihat pada Lampiran 2, pada Gambar 18. Topik utama artikel pada pemberitaan berita online.

4. Kesimpulan dan Saran

Pariwisata di Pulau Lombok berperan penting dalam perkembangan ekonomi, sosial, dan budaya lokal. Destinasi ini menawarkan keindahan alam yang meliputi pantai dan gunung, serta kekayaan budaya suku Sasak. Peningkatan infrastruktur pariwisata telah memperbaiki aksesibilitas bagi wisatawan, menciptakan lapangan pekerjaan, dan memperkuat interaksi sosial antara pengunjung dan penduduk lokal. Analisis model topik dari konten media sosial dan artikel berita menunjukkan adanya empat hingga enam tema utama yang mencakup promosi destinasi, pengembangan desa wisata, dan pengaruh acara olahraga seperti MotoGP terhadap citra pariwisata daerah. Metode pemodelan seperti TF-IDF dan LDA digunakan untuk mengidentifikasi dan mengelompokkan tema-tema tersebut. Hasil pengujian model topik menunjukkan kemampuan prediktif yang kuat dengan nilai perplexity sebesar -9.346, menandakan pemahaman yang baik terhadap distribusi kata. Coherence Score sebesar 0.5873

mencerminkan koherensi tematik yang tinggi, di mana kata-kata dalam setiap topik saling berkaitan secara semantik. Evaluasi model komentar di media sosial menunjukkan perplexity -6.276 dan coherence score 0.513, yang menandakan konsistensi yang baik. Pengujian pada artikel berita online menghasilkan perplexity -8.52, menunjukkan prediksi yang baik.

5. UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih yang sebesar-besarnya kepada Direktorat Riset Teknologi, dan Pengabdian Kepada Masyarakat (DRTPM) Kementerian Pendidikan dan Kebudayaan Republik Indonesia atas dukungan dan fasilitas yang telah diberikan dalam pelaksanaan penelitian ini. Bantuan yang diberikan telah berkontribusi signifikan dalam pengembangan ilmu pengetahuan dan pengabdian kepada masyarakat, terutama dalam konteks pariwisata di Pulau Lombok. Kami berharap hasil penelitian ini dapat memberikan manfaat bagi semua pihak dan mendukung pengambilan keputusan yang lebih baik dalam bidang pariwisata serta pengembangan masyarakat.

DAFTAR PUSTAKA:

- [1] J. Mooris, "Pulau Lombok yang Menakjubkan," Wonderfull Indonesia. Accessed: Jan. 05, 2024. [Online]. Available: <https://www.indonesia.travel/id/id/des-tinasi/bali-nusa-tenggara/lombok>
- [2] P. A. Valguna, D. Dewanti, and L. B. Suparma, "Dampak Perkembangan Pariwisata Pulau Lombok terhadap Pengembangan Bandar Udara Internasional Lombok," *Jurnal Manajemen Aset Infrastruktur & Fasilitas*, vol. 4, no. 3, pp. 195–210, 2020, doi: 10.12962/j26151847.v4i3.7100.
- [3] T. Surahman, I. N. Sudiarta, and I. K. Suwena, "Dampak Pengembangan Pariwisata Terhadap Ekonomi dan Sosial Budaya Masyarakat Lokal Desa Wisata Sasak Ende Kabupaten Lombok Tengah," *Jurnal Analisis Pariwisata*, vol. 20, no. 1, pp. 38–48, 2020, [Online]. Available: <https://ejournal.bsi.ac.id/ejurnal/index.php/khasanah/article/view/9817>
- [4] R. K. Mishra, J. A. A. Jothi, S. Urolagin, and K. Irani, "Knowledge based topic retrieval for recommendations and tourism promotions," *International Journal of Information Management Data Insights*, vol. 3, no. 1, 2023, doi: 10.1016/j.jjime.2022.100145.
- [5] J. H. Lee and M. J. Ostwald, "Latent Dirichlet Allocation (LDA) topic models for Space Syntax studies on spatial experience," *City Territory and Architecture*, vol. 11, no. 1, pp. 1–20, 2024, doi: 10.1186/s40410-023-00223-3.
- [6] H. Laaroussi, F. Guerouate, and M. Sbihi, "Deep learning approach and topic modelling for forecasting tourist arrivals," *International Journal of Electrical and Computer Engineering Systems*, vol. 14, no. 3, pp. 411–419, 2023.
- [7] H. Peng, J. Huang, X. Li, D. Dong, and P. Fan, "Topic Extraction Based on LDA and Its Application in Tourism," in *International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology*, 2022, pp. 52–57.
- [8] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
- [9] M. Kumar and R. Vig, "Term-Frequency Inverse Document Frequency Definition Semantic (TIDS) Based Focused Web Crawler," in *Global Trends in Information Systems and Software Applications*, New York: Springer, 2011, ch. 5, pp. 31–36.
- [10] M. Bakrey, "All about Latent Dirichlet Allocation (LDA) in NLP," Medium. Accessed: Apr. 02, 2024. [Online]. Available: <https://mohamedbakrey094.medium.com/all-about-latent-dirichlet-allocation-lda-in-nlp-6cfa7825034e>
- [11] J. O. Young, "A DEFENCE OF THE COHERENCE THEORY OF TRUTH," 2001.
- [12] F. Jelinek, R. L. Mercer, and L. R. Bahl, "Perplexity—a measure of the difficulty of speech recognition tasks," *J Acoust Soc Am*, vol. 62, no. 1, 2005.
- [13] F. Jelinek, R. L. Mercer, L. R. Bahl, and J. K. Baker, "Perplexity—a measure of the difficulty of speech recognition tasks," *J Acoust Soc Am*, vol. 62, no. S1, pp. S63–S63, Dec. 1977, doi: 10.1121/1.2016299.