



Analisis Topic-Modelling Menggunakan Latent Dirichlet Allocation (LDA) Pada Ulasan Sosial Media Youtube

Vika Alpiana^{1,*}, Abu Salam¹, Farrikh Alzami², Ifan Rizqa¹, Diana Aqmala³

¹Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

²Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Dian Nuswantoro, Semarang, Indonesia

³Fakultas Ekonomi & Bisnis, Program Studi Manajemen, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: ^{1,*}111202012725@mhs.dinus.ac.id, ²abu.salam@dsn.dinus.ac.id, ³alzami@dsn.dinus.ac.id,

⁴rizqa.ifan@dsn.dinus.ac.id, ⁵diana.aqmala@gmail.com

Email Penulis Korespondensi: 111202012725@mhs.dinus.ac.id

Abstrak—Penelitian ini mengeksplorasi peran Usaha Mikro, Kecil, dan Menengah (UMKM) dalam ekonomi Indonesia, fokusnya pada tantangan penjualan dan pemasaran di era media sosial, terutama YouTube. Dengan jutaan individu menggunakan platform ini untuk berbagi pandangan produk, ulasan, dan pengalaman, UMKM perlu mendapatkan umpan balik yang relevan. Studi ini menerapkan text mining, khususnya metode analisis topic modeling dengan Latent Dirichlet Allocation (LDA), untuk menganalisis komentar pengguna pada video UMKM, dengan penekanan pada Lumpia Gang Lombok Semarang di YouTube. Melalui penerapan LDA, identifikasi sepuluh topik utama dilakukan, dengan nilai koherensi tertinggi mencapai 0.414027. Visualisasi intertopic distance map memberikan pemahaman tentang hubungan antar topik dan kata-kata dominan. Analisis komentar memberikan wawasan berharga terkait preferensi dan persepsi pengguna terhadap produk, mendukung UMKM dalam memahami kepuasan pelanggan dan meningkatkan value pada UMKM tersebut. Temuan ini juga menegaskan efektivitas YouTube sebagai sumber data yang relevan untuk memahami preferensi masyarakat terhadap produk UMKM. Penelitian ini merinci metode pengolahan teks, termasuk pengambilan, pembersihan, tokenisasi, normalisasi, penghapusan stopwords, dan stemming. Dengan pendekatan ini, penelitian tidak hanya memberikan wawasan analisis topik dalam konteks media sosial, tetapi juga memberikan kontribusi berharga untuk pengembangan dan pemasaran UMKM melalui pemahaman yang lebih baik terhadap data sosial media, khususnya di platform YouTube.

Kata kunci : Ukm; Analisis Topic Modelling; Latent Dirichlet Allocation (LDA); YouTube; Preferensi Pengguna

Abstract—This research explores the role of Micro, Small, and Medium Enterprises (MSMEs) in the Indonesian economy, focusing on sales and marketing challenges in the era of social media, especially YouTube. With millions of individuals using this platform to share product insights, reviews, and experiences, MSMEs need to receive relevant feedback. This study applies text mining, particularly the topic modeling analysis method with Latent Dirichlet Allocation (LDA), to analyze user comments on MSME videos, with an emphasis on Lumpia Gang Lombok Semarang on YouTube. Through the application of LDA, the identification of ten main topics is conducted, with the highest coherence value reaching 0.414027. The visualization of the intertopic distance map provides an understanding of the relationships between topics and dominant words. Comment analysis provides valuable insights into user preferences and perceptions of products, supporting MSMEs in understanding customer satisfaction and enhancing value for those enterprises. These findings also affirm the effectiveness of YouTube as a relevant data source for understanding public preferences for MSME products. This research details text processing methods, including extraction, cleaning, tokenization, normalization, removal of stopwords, and stemming. With this approach, the research not only provides insights into topic analysis in the context of social media but also makes a valuable contribution to the development and marketing of MSMEs through a better understanding of social media data, especially on the YouTube platform.

Keywords: MSMEs; Topic Modeling Analysis; Latent Dirichlet Allocation (LDA); YouTube; User Preferences

1. PENDAHULUAN

UMKM atau usaha mikro kecil menengah merupakan sektor penting dalam perekonomian Indonesia karena berkontribusi terhadap pembangunan nasional dan menciptakan lapangan kerja [1]. Jumlah Usaha Mikro, Kecil, dan Menengah (UMKM) di Indonesia saat ini mencapai 64,2 juta, dan kontribusinya terhadap Produk Domestik Bruto (PDB) mencapai 60,6% [2]. Meskipun memiliki peran krusial dalam perekonomian, UMKM sering menghadapi sejumlah tantangan terutama terkait dengan aspek penjualan seperti kurangnya pengetahuan tentang pemasaran produk melalui branding dan penggunaan teknologi [3].

Dalam era digital yang sedang berkembang pesat dan tren media sosial yang semakin meluas, tingginya minat masyarakat terhadap youtube telah menjadikan platform media sosial ini semakin digunakan secara luas oleh berbagai segmen masyarakat [4]. YouTube menduduki peringkat pertama sebagai platform media sosial paling populer di Indonesia dan menghasilkan volume data yang besar dalam bentuk komentar [5]. Perubahan terjadi pada lingkungan bisnis dan konsumen sejak media sosial youtube menjadi tempat berbagi ulasan dan tanggapan dari pelanggan. Jutaan individu aktif memanfaatkan platform media sosial youtube untuk berbagi pandangan, opini, dan pengalaman mereka terkait produk, layanan, dan merek tertentu. Ulasan ini tidak hanya menjadi sumber informasi berharga secara kuantitatif dan kualitatif bagi bisnis, tetapi juga memperlihatkan dampak signifikan media sosial pada dinamika interaksi antara pelanggan dan Perusahaan [6]. Oleh karena itu, UMKM memerlukan umpan balik mengenai produk mereka untuk memahami topik yang sedang dibahas, maka diperlukan penerapan text mining. Dengan mengetahui topik ini, produsen dapat meningkatkan value pada UMKM tersebut. Oleh karena itu, YouTube menjadi pilihan yang efisien untuk pengambilan data komentar. Dengan



menganalisis komentar-komentar ini, UMKM dapat memahami kepuasan pelanggan, mengidentifikasi area perbaikan, dan menyesuaikan penawaran untuk memenuhi kebutuhan pelanggan dengan lebih baik. Salah satu video yang diunggah oleh pengguna ke dalam platform YouTube ialah lumpia gang Lombok Semarang yang merupakan salah satu UMKM yang ada di Jawa Tengah.

Penelitian ini akan mengimplementasikan metode analisis topic modelling terhadap komentar di media sosial youtube. Definisi dari pemodelan topik adalah suatu model generatif dokumen yang menerapkan suatu prosedur probabilistik untuk menghasilkan dokumen, di mana hasil dari pemodelan topik adalah kelompok kata yang membentuk cluster dan topik berdasarkan distribusi probabilitas kata dalam suatu kumpulan dokumen [7]. Penelitian ini menggunakan teknik pengolahan teks dan menggunakan pemodelan topik dengan Latent Dirichlet Allocation (LDA) [8]. Dengan menggunakan LDA pada komentar dan ulasan dari media sosial Youtube, peneliti dapat mengungkap tema dan subjek mendasar yang dibahas oleh pengguna [9]. metode ini berguna untuk mengidentifikasi hubungan tersembunyi di antara dokumen teks dan mengungkap topik yang ada di dalamnya [10].

Berdasarkan penelitian yang dilakukan oleh K. Amaradiena dan T. Widarmanti [11], dengan judul "LDA-Topic Modeling: Menggunakan Ulasan Pengguna Untuk Meningkatkan User Experience (Studi pada PeduliLindungi)" ditemukan bahwa penggunaan LDA-Topic Modeling untuk menganalisis ulasan pengguna terhadap aplikasi Peduli Lindungi menghasilkan 9 topik dengan nilai coherence 0,4381, dengan topik yang paling umum terkait dengan penggunaan aplikasi. Selain itu, penelitian ini juga mengidentifikasi 4 metrik UX: kebahagiaan, keterlibatan, retensi, dan keberhasilan tugas. Temuan ini memberikan wawasan yang berharga terkait persepsi pengguna dan pengalaman mereka dengan aplikasi PeduliLindungi, serta menunjukkan pentingnya pengembangan dan perbaikan berkelanjutan untuk memastikan fungsionalitas optimal dan kepuasan pengguna. Dengan demikian, penelitian ini menunjukkan bahwa LDA adalah alat yang efektif untuk analisis topik dalam konteks ulasan pengguna, dengan nilai koherensi yang tinggi sehingga model yang dihasilkan dapat diinterpretasikan dengan baik oleh manusia.

Penelitian yang dilakukan oleh Ni Luh Putu Merawati Putu, Ahmad Zuli Amrullah, dan Ismarmiaty [12], dengan judul "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation" fokus pada identifikasi topik yang paling umum dibahas oleh wisatawan dalam ulasan mereka, seperti keindahan alam, akomodasi, dan kuliner. Dapat disimpulkan bahwa hasil penelitian ini memiliki potensi untuk memberikan kontribusi yang berharga bagi pengambil keputusan, termasuk pemerintah daerah, pelaku wisata, dan stakeholder, dalam upaya pengembangan dan pengelolaan pariwisata di Lombok. Tabel 5 dalam jurnal tersebut menunjukkan nilai koherensi untuk berbagai topik dalam kategori positif dan negatif. Sebagai contoh, nilai koherensi tertinggi untuk kategori positif dicapai oleh kelompok ke-8 dengan skor 0,613, sementara untuk kategori negatif dicapai oleh kelompok ke-12 dengan skor 0,528. Semakin tinggi tingkat koherensi dalam suatu topik, semakin mudah untuk menginterpretasikan topik tersebut berdasarkan sekumpulan kata yang membentuknya. Oleh karena itu, nilai koherensi ini memberikan gambaran tentang tingkat kemudahan interpretasi semantik suatu topik.

Penelitian yang dilaksanakan oleh Diandra Zakeshia Tiara Kannitha, Mustafid Mustafid, dan Puspita Kartikasari [13], dengan judul "Pemodelan Topik Pada Keluhan Pelanggan Menggunakan Algoritma Latent Dirichlet Allocation Dalam Media Sosial Twitter" mengindikasikan bahwa keluhan pelanggan terhadap layanan internet dari First Media dan IndiHome berfokus pada masalah seperti masalah internet yang mati, putus, dan lambat. Algoritma Latent Dirichlet Allocation dipilih untuk penelitian ini karena kinerjanya yang terbukti baik. Berdasarkan hasil interpretasi, terungkap bahwa 70% untuk First Media dan 81,81% untuk IndiHome dari topik ini sesuai dengan keluhan pelanggan yang diungkapkan melalui tweet mereka. Informasi dari topik yang telah diidentifikasi dapat dijadikan sebagai bahan evaluasi bagi perusahaan guna menjaga kualitas layanan dan kesetiaan pelanggan.

Penelitian yang dilakukan Musliadi K H, Hazriani Zainuddin, dan Yuyun [14], dengan judul "Twitter Social Media Conversion Topic Trending Analysis Using Latent Dirichlet Allocation Algorithm" mengeksplorasi penggunaan algoritma Latent Dirichlet Allocation (LDA) untuk memodelkan dan memvisualisasikan topik yang dibicarakan di media sosial oleh komunitas Makassar di Twitter. Studi ini menemukan bahwa topik kedelapan adalah kata yang paling sering digunakan dalam percakapan, sedangkan topik ke-7 dan ke-6 muncul sebagai inti percakapan berdasarkan penyebaran kata-kata dengan frekuensi istilah tertinggi. kapitalisasi dan visualisasi menggunakan metode LDA menghasilkan kata-kata dengan tren tertinggi dan topik dengan frekuensi istilah tertinggi.

Penelitian yang dilakukan Muhammad Habibi, Adri Priadana, Andika Bayu Saputra, dan Puji Winar Cahyo [15], dengan judul "Topic Modelling of Germas Related Content on Instagram Using Latent Dirichlet Allocation (LDA)" bertujuan untuk memodelkan topik pada caption Instagram terkait Gerakan Masyarakat Hidup Sehat (GERMAS) dengan menggunakan teknik Latent Dirichlet Allocation (LDA). Studi ini menggunakan koherensi topik untuk mengevaluasi jumlah topik yang paling sesuai, sehingga menghasilkan delapan segmen topik Analisis isi pada setiap segmen topik menunjukkan bahwa topik terkait GERMAS yang paling dominan adalah pola makan gaya hidup sehat.

Berdasarkan latar belakang tersebut, tujuan dari penelitian ini adalah menerapkan metode analisis pemodelan topik, terutama dengan menggunakan Latent Dirichlet Allocation (LDA), pada komentar-komentar

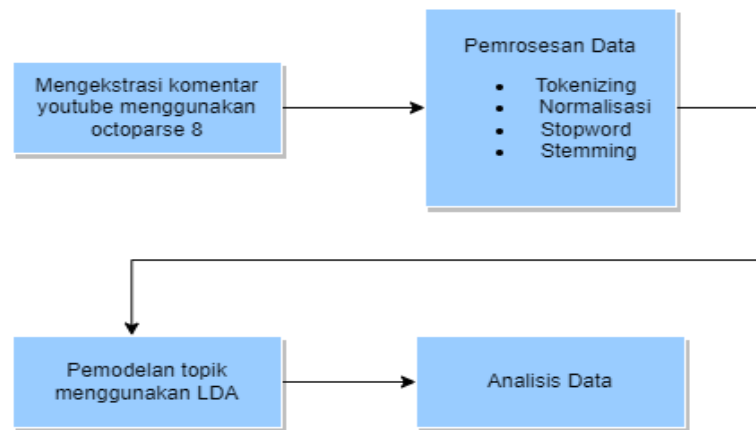


yang ada di platform media sosial YouTube. Saat ini, belum ada penelitian yang menggunakan data dari YouTube untuk melakukan pemodelan topik terkait Usaha Mikro Kecil Menengah (UMKM), khususnya dalam konteks Lumpia Gang Lombok Semarang. Oleh karena itu, penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang persepsi pelanggan terhadap UMKM Lumpia Gang Lombok Semarang, serta berkontribusi pada pengembangan metode analisis topik dalam konteks media sosial.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahap-tahapan penelitian telah disusun sesuai dengan petunjuk pada gambar 1. Berikut adalah penjelasan mengenai serangkaian langkah penelitian yang dijalankan guna mencapai tujuan penelitian tersebut. Metodologi penelitian kami melibatkan beberapa tahapan: pertama, kami mengumpulkan data komentar dari YouTube. Kemudian, kami melakukan pembersihan data, tokenisasi, normalisasi, stopword, stemming, dan tahap terakhir melibatkan pemodelan topik menggunakan LDA, diikuti oleh proses Analisis Data. Oleh karena itu, pada gambar 1 adalah gambar yang menunjukkan tahapan-tahapan penelitian yang akan dilakukan untuk melaksanakan proses penelitian ini.



Gambar 1. Diagram Metode Penelitian

2.2 Pengambilan Data

Tahap awal dalam melakukan topic modelling adalah pengumpulan data. Sumber data yang dianalisis dalam penelitian ini berasal langsung dari tiga akun YouTube dengan menggunakan kata kunci "Lumpia Gang Lombok Semarang". Proses pengambilan data dilakukan menggunakan Octoparse 8, dengan dukungan Google Colab, dan menggunakan bahasa pemrograman Python.

2.3 Pembersihan Data (preprocessing)

Preprocessing bertujuan mengubah data dari text yang semula tidak terstruktur menjadi data yang tersusun secara terstruktur. Proses ini dimaksudkan untuk menghilangkan gangguan, menyempurnakan bentuk teks, dan mengurangi jumlah kata [13]. Selain itu, karena data yang diproses berasal dari ulasan media sosial, maka perlu penanganan khusus seperti menghapus URL dan alamat email, menghapus emoticon, menghapus media gambar, video dan menghapus bilangan angka. Preprocessing memiliki dampak yang signifikan terhadap hasil analisis dan akurasi keakuratan dalam text mining [14]. Berikut merupakan contoh hasil dari pembersihan data ditunjukkan pada tabel 1 dibawah ini:

Tabel 1. Hasil Pembersihan Data

Text	Hasil Pembersihan Text
Jadi LuMpia atau LuNpia?? 😊😊	jadi lumpia atau lulpia
HALAL ya Koh??	halal ya koh
Koko Koko yang suka pedas👍👍	koko koko yang suka pedas
uh dapurnya udah bagus ya kaya standar rest	uh dapurnya udah bagus ya kaya standar resto
HM yamy 😊😊😊😊	hm yamy

2.4 Tokenisasi

Tokenisasi adalah proses memisahkan teks menjadi kata-kata Tunggal [15]. proses ini menggunakan pemrosesan teks menggunakan library Natural Language Toolkit (NLTK) dalam konteks tokenisasi kata. Berikut merupakan contoh hasil dari tokenisasi ditunjukkan pada tabel 2 dibawah ini:



Tabel 2. Hasil Tokenisasi

Processed Text	Tokens
jadi lumpia atau lumpia	[jadi, lumpia, atau, lumpia]
halal ya koh	[halal, ya, koh]
koko koko yang suka pedas	[koko, koko, yang, suka, pedas]
isi lupianya mantab ukuranya gede	[isi, lupianya, mantab, ukuranya, gede]

2.5 Normalisasi

Normalisasi adalah suatu proses di mana kata-kata yang tidak umum atau abnormal diubah menjadi bentuk kata yang normal atau standar [16]. Proses ini menggunakan kamus perbaikan kata untuk mendapatkan data yang lebih baik dan secara konsisten meningkatkan kinerja system. Berikut merupakan contoh hasil dari normalisasi ditunjukkan pada tabel 3 dibawah ini:

Tabel 3. Hasil Normalisasi

Processed Text	Hasil Normalisasi
HALAL ya Koh??	halal iya koh
Isi Rebung , yg maling joss	isi rebung maling bagus
Padaahal digulung asal2an koq bisa rapi ya...	gulung asal2an koq rapi iya

2.6 Stopword

Stopword adalah langkah untuk menghapus kata-kata yang kurang bermakna atau merupakan kata sambung [17]. Dalam proses ini menggunakan daftar stopwords bahasa Indonesia dari library Natural Language Toolkit (NLTK) untuk menghapus kata-kata tersebut. Berikut merupakan contoh hasil dari stopwords ditunjukkan pada tabel 4 dibawah ini:

Tabel 4. Hasil Stopword

Processed Text	Hasil Stopword
Mantaapppp cuyyyyy.....	mantap cuy
asli sih lumpia makanan favorit bgt, tapi yang di goreng	asli sih lumpia makan favorit banget goreng
Ada harga ada kualitas	harga kualitas

2.7 Stemming

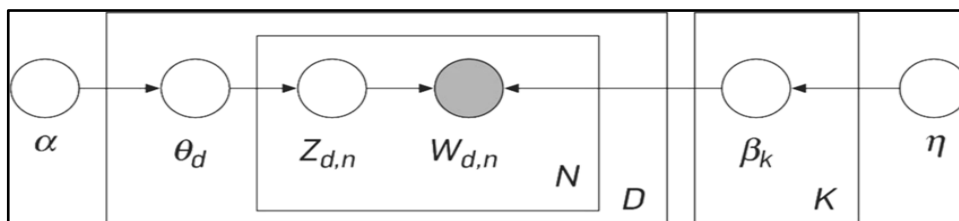
Pada proses ini bertujuan untuk menghilangkan kata yang berimbuhan sehingga menjadi kata dasar [15]. Dengan tujuan mengurangi gangguan dan jumlah dimensi pada data [18]. Dalam tahap ini, pemanfaatan library Sastrawi untuk melakukan proses stemming telah terbukti efektif dalam mengatasi permasalahan over-stemming dan under-stemming, sekaligus meningkatkan efisiensi waktu pemrosesan [19]. Berikut merupakan contoh hasil dari stemming ditunjukkan pada tabel 5 dibawah ini:

Tabel 5. Hasil Stemming

Processed Text	Stemming
Lumpia basah the best, makan 1 meti nagih heheheee	lumpia basah the best makan meti nagih hehehe
Tapi resiko bau rebung..	resiko bau rebung
Bikin ngiler aja koh...	ngiler koh

2.8 Pemodelan Topik menggunakan LDA

Dalam fase penelitian ini, dilakukan pemodelan topik dengan menggunakan metode LDA. Pemodelan topik dapat diartikan sebagai suatu algoritma yang digunakan untuk menemukan tema utama dari sekelompok dokumen yang besar dan tidak terorganisir [15]. Latent Dirichlet Allocation (LDA) adalah salah satu metode pembelajaran tanpa pengawasan probabilitas yang paling stabil dalam menemukan struktur pola yang tersembunyi [18].



Gambar 2. Representasi Model LDA (Latent Dirichlet Allocation)

Setiap topik pada gilirannya terdiri dari distribusi kata-kata yang mewakili topik tersebut. Pendekatan ini mengasumsikan bahwa dokumen-dokumen dalam koleksi bersifat probabilistik, dimana setiap kata dalam



dokumen dapat dijelaskan oleh salah satu dari topik yang ada [12]. Proses ini dapat dianalogikan sebagai sebuah model probabilistic, seperti yang diilustrasikan dalam gambar diatas. Berdasarkan ilustrasi pada gambar 2 maka LDA dapat digambarkan khusus penggunaan notasi- notasi matematika yang dapat dilihat pada persamaan berikut

$$p(\beta_{1:k} \theta_{1:D} Z_{1:D} W_{1:D}) = \prod_{i=1}^k p(\beta_i) \prod_{d=1}^D p(\theta_d) \tag{1}$$

$$\prod_{n=1}^N p(Z_{d,n} | \theta_d) p(W_{d,n} | \beta_{1:k} Z_{d,n}) \tag{2}$$

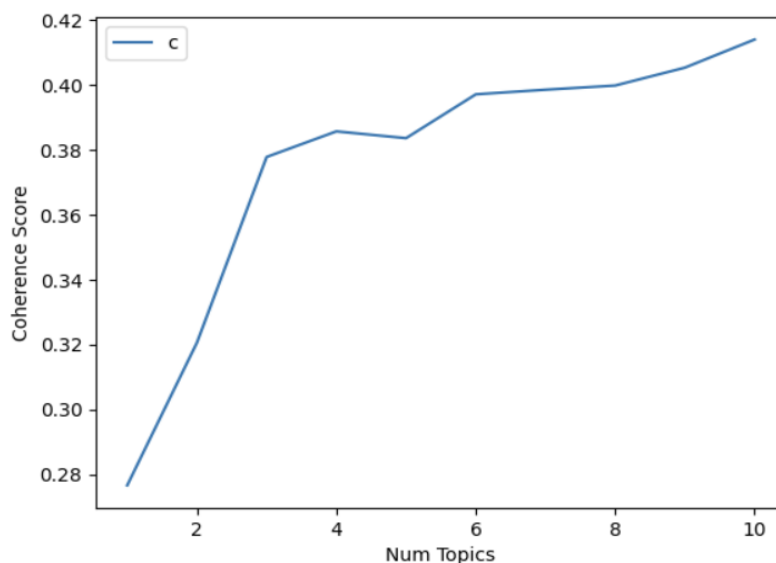
D adalah himpunan dokumen, K adalah himpunan topik, N adalah jumlah kata dalam dokumen diberi indeks (N_d), W_(d,n) adalah kata ke-n dalam dokumen d, Z_(d,n) adalah topik ke-n dalam dokumen d, θ_d adalah distribusi topik dalam dokumen d, β_k adalah distribusi kata dalam topik k pada kosakata, serta α, η adalah parameter dirchlect [12].

2.9 Analisis Data

Dalam langkah ini, analisis kualitatif dilakukan dengan memperhatikan istilah-istilah yang muncul setelah topik dikelompokkan. Informasi yang terkandung dalam pengelompokan topik tersebut dianalisis dengan tujuan memahami makna yang terkait dengan data.

3. HASIL DAN PEMBAHASAN

Proses pengumpulan data mengenai Lumpia Gang Lombok Semarang melibatkan tiga sumber utama dari saluran YouTube yang membahas masakan tersebut. Pertama-tama, "Goyang Laper" memberikan kontribusi sebanyak 86 komentar yang menjadi objek analisis. Sumber kedua berasal dari saluran "Kubiler," yang menyediakan data sebanyak 235 komentar untuk dianalisis. Selanjutnya, "Anak Kuliner" menjadi sumber ketiga dengan menyumbangkan 203 data komentar. Dengan menggabungkan ketiga sumber ini, diperoleh total 524 komentar yang akan menjadi dasar untuk analisis lebih lanjut. Teknik pengumpulan data dilakukan dengan menggunakan Octoparse 8, memastikan keakuratan dan konsistensi informasi yang diperoleh dari ketiga saluran YouTube. Setelah berhasil mengumpulkan data, dilakukan tahapan preprocessing untuk membersihkan data dari noise atau gangguan yang mungkin muncul selama proses pengumpulan. Preprocessing adalah langkah penting untuk memastikan data yang digunakan dalam analisis memiliki kualitas yang baik dan dapat diandalkan. Sebagai langkah selanjutnya, nilai koherensi diaplikasikan sebagai metrik untuk mengevaluasi sejauh mana keakuratan topik dari dokumen yang terkumpul [20]. Tingkat koherensi yang tinggi memberikan indikasi bahwa hasil interpretasi dalam pemodelan topik lebih baik [21]. Berdasarkan gambar 4 yang terlampir di bawah, dapat diamati bahwa setiap topik yang dihasilkan menunjukkan perbedaan dalam nilai koherensi.



Gambar 3. Grafik Garis Nilai Koherensi

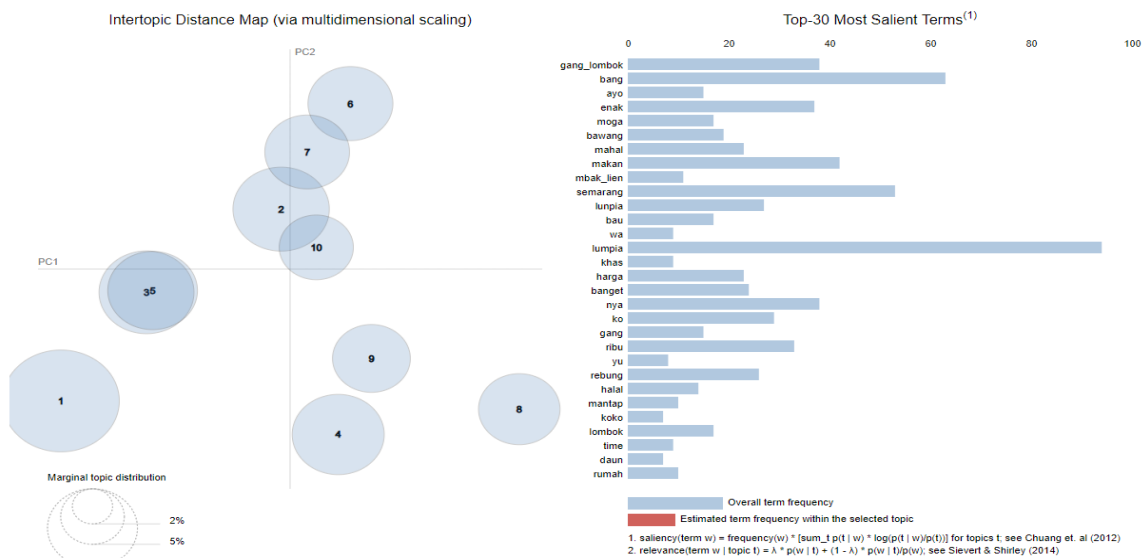
Dalam data yang terdokumentasi dalam tabel di bawah, dapat diperhatikan bahwa nilai koherensi tertinggi terdeteksi pada topik ke-10, dengan nilai mencapai 0.414027. Angka ini mengindikasikan bahwa dari keseluruhan sepuluh topik yang dihasilkan, topik ke-10 menunjukkan tingkat konsistensi dan keterkaitan yang paling optimal dengan konten dokumen yang sedang dianalisis. Dengan mengeksplorasi nilai koherensi pada setiap topik, dapat ditemukan perincian lebih lanjut mengenai sejauh mana interpretasi dan keakuratan pemodelan topik pada data komentar yang telah dikumpulkan. Informasi lebih lanjut terkait nilai koherensi untuk setiap topik dapat diakses dan dieksplorasi secara lebih mendalam melalui tabel 6 yang disajikan di bawah ini.



Tabel 6. Hasil Nilai Koherensi Pada Setiap Topik

Topik	Nilai Koherensi
1	0.276591
2	0.320587
3	0.377822
4	0.385708
5	0.383615
6	0.397127
7	0.398551
8	0.399841
9	0.405358
10	0.414027

Setelah memperoleh nilai koherensi untuk setiap topik, hasilnya disajikan dalam bentuk visualisasi intertopic distance map, yang dapat dilihat pada gambar 5. Visualisasi ini memberikan informasi terkait dengan topik-topik yang ada dan kata-kata yang paling mendominasi masing-masing topik. Peta jarak antartopik ini membantu dalam memahami hubungan dan struktur antar topik, serta mengidentifikasi kata-kata kunci yang secara signifikan berkontribusi pada setiap topik dalam pemodelan topik menggunakan LDA.



Gambar 3. Visualisasi Intertopic Distance Map

Ilustrasi di atas memberikan gambaran tentang 30 dominan yang muncul dalam korpus, dengan pembagian menjadi dua bagian, yaitu panel kiri dan panel kanan. Pada panel kiri, terdapat representasi jarak antar topik yang mengindikasikan bahwa Topik 3 dan Topik 5 saling berdekatan, menunjukkan adanya kesamaan dan kedekatan dalam pembahasan kedua topik tersebut. Sementara pada Topik 10, 2, 7, dan 6, terlihat kesamaan dan keterkaitan antar keempat topik tersebut. Adapun untuk Topik 1, 4, 8, dan 9, jaraknya jauh dari topik-topik lainnya, menandakan bahwa topik-topik tersebut tidak memiliki kedekatan dan kesamaan. Di sisi lain, pada panel kanan, perhatian difokuskan pada kata-kata yang dominan dan sering muncul dalam korpus, seperti "lumpia," "bang," "Semarang," "makan," "gang_lombok," "enak," dan sebagainya. Keberadaan kata-kata ini menyoroti aspek-aspek yang paling sering diakses atau diangkat dalam konteks korpus, memberikan pandangan lebih mendalam terhadap inti dari informasi yang terkandung di dalamnya. Penafsiran makna dari suatu topik dilihat dari berdasarkan 30 kata yang paling dominan seperti yang ditunjukkan pada Tabel 7.

Tabel 7. Hasil 30 Kata Dominan

Index	Kata	Hasil
0	Lumpia	Lumpia mendapat perhatian khusus dalam ulasan di media sosial YouTube, menjadi fokus utama dalam menilai kualitas, rasa, dan presentasi kuliner dari UMKM yang diulas..
1	Bang	"Panggilan akrab 'Bang' digunakan secara informal di beberapa daerah di Indonesia dan dapat menciptakan hubungan personal antara reviewer dan pemilik UMKM dalam review YouTube, menunjukkan kedekatan dan keakraban."



Index	Kata	Hasil
2	Semarang	Lokasi Lumpia Gang Lombok dapat diidentifikasi di Semarang berdasarkan letak geografisnya.
3	Makan	Dalam konteks review, kata "Makan" menyoroti pengalaman konsumen serta aspek rasa dan kualitas hidangan dari UMKM.
4	Nya	Dalam review, kata ini dapat menunjukkan hubungan antara subjek dan objek, seperti "rasa lumpianya" atau "harga makanannya".
5	Gang_Lombok	"Gang Lombok" dalam review dapat menjadi faktor penilaian lokasi UMKM, memberikan informasi tambahan tentang tempat atau kawasan usaha tersebut kepada penontonton yang lain.
6	Enak	"Enak" mengindikasikan kesan positif terhadap rasa makanan, digunakan dalam review YouTube untuk merekomendasikan makanan tertentu kepada penonton yang lain.
7	Ribu	"Ribu" merujuk pada satuan angka ribu dan dalam konteks review YouTube, kata ini menyoroti informasi harga produk atau makanan, memberikan pandangan mengenai kisaran harga yang diantisipasi kepada penonton.
8	Ko	"Ko" bisa digunakan sebagai panggilan mirip "Bang" dalam konteks review, untuk menciptakan hubungan personal dan keakraban.
9	Lunpia	"Lunpia" berasal dari dialek bahasa Cina Hokkian "run bing" dan bersinergi dengan bahasa Jawa "lum ping" (kulit), menjadi "lumpia". Ini memberikan konteks sejarah dan asal-usul kata, menambah dimensi keberagaman budaya dalam review.
10	Rebung	Merujuk pada bahan makanan sebagai isian dalam lumpia.
11	Banget	Dalam review, digunakan untuk menyoroti aspek positif atau negatif dari pengalaman makan di Lumpia Gang Lombok Semarang.
12	Harga	"Harga" merujuk pada nilai moneter terkait lumpia atau makanan. Dalam review YouTube, ini memberikan informasi kisaran harga untuk membantu pemirsa dalam membuat keputusan pembelian.
13	Mahal	Mengindikasikan sesuatu yang mahal, terutama dalam konteks makanan seperti lumpia. Poin ini kritis dalam review, memberikan pandangan apakah harga lumpia sebanding dengan kualitas dan pengalaman yang ditawarkan oleh UMKM.
14	Bawang	"Bawang" merujuk pada daun bawang yang digunakan sebagai lalapan pelengkap makan lumpia.
15	Bau	Digunakan untuk menggambarkan aroma Lumpia Gang Lombok Semarang.
16	Moga	Merujuk pada harapan atau doa yang menciptakan atmosfer positif dalam review YouTube, mengekspresikan optimisme bahwa UMKM yang diulas akan berhasil dan disukai oleh konsumen.
17	Lombok	Dalam review, ini merujuk pada lokasi bisnis UMKM yang berada di Gang Lombok Semarang.
18	Gang	Merujuk pada lokasi bisnis UMKM yang berada di Gang Lombok Semarang.
19	Ayo	Merupakan singkatan dari "ayok," ajakan untuk mencoba lumpia di Gang Lombok, Semarang.
20	Halal	Merujuk pada status halal makanan yang diulas, aspek penting bagi konsumen Muslim.
21	Mbak_lien	Merujuk pada identitas pemilik UMKM di lokasi lain.
22	Rumah	Dalam review, merujuk pada kenyamanan tempat UMKM yang diulas.
23	Mantap	Untuk menyampaikan kepuasan atau persetujuan terhadap Lumpia Gang Lombok Semarang.
24	Khas	Merujuk pada elemen unik dari Lumpia Gang Lombok yang membedakannya dari lumpia lainnya.
25	Wa	Wa adalah singkatan dari WhatsApp yang sering digunakan untuk berkomunikasi atau berinteraksi dengan pemilik umkm. Dalam review ini merujuk pada cara pembeli berkomunikasi atau berinteraksi dengan pihak umkm
26	Time	Hal ini berkaitan dengan waktu operasional dari lumpia gang lombok semarang
27	Daun	Merujuk ke bahan yang digunakan atau rasa spesifik dalam lumpia gang lombok semarang yang sedang direview
28	Yu	Merupakan singkatan dari kata "yuk" atau ajakan untuk bergabung atau mencoba lumpia gang lombok semarang
29	koko	"Koko" adalah sebutan untuk kakak laki-laki dalam keluarga Tionghoa-Indonesia. Dalam konteks review, ini bisa digunakan untuk menciptakan hubungan personal dan keakraban dengan pemilik umkm ataupun konten kreator



Melalui analisis ini, dapat terlihat bahwa kata-kata dominan memberikan wawasan tambahan untuk memahami berbagai aspek yang dianalisis. Dalam pemodelan topik menggunakan LDA, setiap kata diberikan nilai bobot yang mencerminkan tingkat kontribusinya terhadap suatu topik tertentu. Proses ini melibatkan analisis statistik terhadap kumpulan dokumen atau teks untuk menentukan pola kemunculan kata-kata tertentu dalam konteks tertentu. Bobot ini memberikan indikasi seberapa penting atau signifikan suatu kata dalam merepresentasikan esensi dari suatu topik. Bobot tiap kata dapat dilihat pada Tabel 8 dibawah ini.

Tabel 8. Hasil Nilai Bobot Tiap Kata

Topic	Words
0	0.030*"semarang" + 0.028*"bang" + 0.025*"lumpia" + 0.020*"iya" + 0.016*"bawang" + 0.015*"ribu" + 0.009*"merah" + 0.009*"daun" + 0.009*"koko" + 0.009*"makan"
1	0.028*"moga" + 0.022*"lumpia" + 0.018*"ko" + 0.012*"iya" + 0.012*"bau" + 0.010*"rebung" + 0.010*"mantap" + 0.007*"ngiler" + 0.007*"umur" + 0.007*"jual"
2	0.051*"lumpia" + 0.028*"semarang" + 0.019*"gang_lombok" + 0.017*"makan" + 0.015*"bang" + 0.011*"iya" + 0.011*"banget" + 0.010*"gang" + 0.010*"kota" + 0.009*"etnis"
3	0.038*"bang" + 0.032*"enak" + 0.025*"makan" + 0.017*"iya" + 0.016*"banget" + 0.014*"mahal" + 0.012*"lumpia" + 0.011*"ramadan" + 0.011*"ko" + 0.011*"sih"
4	0.029*"lumpia" + 0.018*"mbak_lien" + 0.016*"bang" + 0.016*"ayo" + 0.015*"rebung" + 0.015*"mahal" + 0.014*"ribu" + 0.014*"ko" + 0.013*"lumpia" + 0.013*"harga"
5	0.021*"bang" + 0.018*"iya" + 0.016*"enak" + 0.015*"khas" + 0.014*"lumpia" + 0.014*"nya" + 0.012*"bawang" + 0.012*"semarang" + 0.008*"yu" + 0.008*"jawa"
6	0.039*"gang_lombok" + 0.026*"ribu" + 0.026*"semarang" + 0.024*"lumpia" + 0.022*"nya" + 0.021*"enak" + 0.018*"lumpia" + 0.017*"lombok" + 0.017*"harga" + 0.014*"halal"
7	0.028*"lumpia" + 0.021*"makan" + 0.020*"nya" + 0.017*"wa" + 0.017*"banget" + 0.017*"bawang" + 0.014*"koh" + 0.012*"enak" + 0.011*"air" + 0.010*"semarang"
8	0.020*"mahal" + 0.017*"bau" + 0.015*"rebung" + 0.015*"beli" + 0.013*"lumpia" + 0.013*"harga" + 0.013*"ngiler" + 0.011*"beda" + 0.010*"ko" + 0.010*"time"
9	0.049*"lumpia" + 0.043*"bang" + 0.023*"makan" + 0.020*"ayo" + 0.017*"semarang" + 0.016*"nya" + 0.015*"iya" + 0.012*"rumah" + 0.011*"rebung" + 0.008*"suka"

Setelah menyelesaikan proses pemodelan topik menggunakan Latent Dirichlet Allocation (LDA) pada media sosial YouTube, hasil topik dan penjelasannya dapat dilihat pada Tabel 9 di bawah ini.

Tabel 9. Hasil Analisis 10 Topik

Index	Topik	Hasil
0	Bahan baku	Hal ini berkaitan dengan Kualitas rasa produk atau pengalaman makan konsumen sangat terkait dengan jenis dan mutu bahan baku yang digunakan.
1	Cita rasa	Cita rasa memiliki peran yang sangat penting dalam dunia kuliner, sehingga kemajuan lumpia Gang Lombok Semarang dapat dicapai dengan baik apabila cita rasanya memuaskan.
2	Lokasi	Lokasi geografis Lumpia Gang Lombok Semarang memainkan peran penting dalam daya tarik dan keterjangkauan bagi konsumen. Keistimewaan lokasinya menjadi faktor utama yang membuat Lumpia Gang Lombok Semarang menjadi destinasi kuliner yang menarik.
3	Ramadhan	Keterkaitan lumpia gang lombok semarang dengan bulan Ramadhan mencakup bagaimana lumpia ini menjadi pilihan makanan yang populer selama bulan puasa dan bagaimana lumpia gang lombok ini bisa menyesuaikan dengan suasana puasa.
4	Rekomendasi tempat lain	Penonton memberikan rekomendasi tempat lain yang sebanding atau bahkan lebih baik. Hal ini dapat memberikan perspektif tambahan tentang sejauh mana Lumpia Gang Lombok Semarang bersaing dengan tempat makan serupa.
5	Ciri Khas	Elemen khusus atau ciri khas dari Lumpia Gang Lombok Semarang bisa melibatkan metode penyajian yang unik atau penambahan bahan tertentu yang membuatnya berbeda dari jenis lumpia lainnya.
6	Kehalalan	Aspek kehalalan produk yang dijual di Lumpia Gang Lombok Semarang dapat berpengaruh pada minat konsumen, terutama bagi mereka yang memperhatikan aspek keagamaan dalam pemilihan makanan.
7	Pengalaman makan	Pengalaman makan di Lumpia Gang Lombok Semarang dapat membantu mengidentifikasi elemen-elemen tertentu, seperti pelayanan, kebersihan, atau kecepatan pelayanan. Pengaruhnya dapat terlihat dalam citra dan reputasi Lumpia Gang Lombok Semarang di mata konsumen.
8	Harga	Harga dapat memberikan wawasan tentang bagaimana konsumen menilai hubungan yang erat antara kualitas dan harga di Lumpia Gang Lombok Semarang. Hal ini



Index	Topik	Hasil
9	Ajakan	dapat terlihat dalam daya tarik Lumpia Gang Lombok Semarang terhadap segmen pasar tertentu, serta sejauh mana harga memainkan peran dalam keputusan konsumen untuk mengunjungi tempat tersebut. Ajakan atau rekomendasi dari para pengguna YouTube untuk mencoba Lumpia Gang Lombok Semarang dapat meningkatkan potensi pertumbuhan jumlah pengunjung atau pelanggan baru.

Dalam analisis ulasan media sosial YouTube mengenai Lumpia Gang Lombok Semarang, terdapat korelasi yang signifikan antara kata-kata dominan dan topik yang telah dianalisis menggunakan LDA. Poin utama, seperti "Lumpia" dan "Rebung," erat terkait dengan topik bahan baku (Topik 1), menunjukkan bahwa konsumen memberikan perhatian khusus pada jenis dan mutu bahan baku dalam menilai rasa produk. Selain itu, kata "Enak" terhubung dengan topik cita rasa (Topik 2), mengindikasikan kepuasan konsumen terhadap pengalaman rasa lumpia tersebut. Lokasi Lumpia Gang Lombok Semarang, terutama dengan kata-kata seperti "Gang Lombok" dan "Semarang," berkorelasi dengan topik lokasi (Topik 3), menunjukkan bahwa lokasi usaha memainkan peran kunci dalam penilaian dan daya tarik konsumen. Korelasi lain terlihat dalam kata "Ramadhan," yang terhubung dengan topik keterkaitan Lumpia Gang Lombok Semarang dengan bulan puasa (Topik 4), mencerminkan popularitas lumpia sebagai pilihan makanan selama bulan Ramadhan. Rekomendasi tempat lain (Topik 5) juga muncul, menunjukkan bahwa penonton memberikan perbandingan antara Lumpia Gang Lombok Semarang dengan tempat makan serupa. Selain itu, kata "Ciri Khas" berkorelasi dengan topik ciri khas (Topik 6), menunjukkan bahwa konsumen mungkin menyoroti keunikan atau elemen khusus dari Lumpia Gang Lombok Semarang. Aspek kehalalan (Topik 7) tercermin melalui kata "Halal," dan pengalaman makan (Topik 8) dibahas dengan kata "Pengalaman makan" dan "Mantap," menunjukkan bahwa konsumen memberikan umpan balik terkait pelayanan dan kualitas. Harga (Topik 9) juga menjadi sorotan, dengan kata "Harga" memberikan wawasan tentang bagaimana konsumen menilai hubungan antara kualitas dan harga. Terakhir, kata "Ajakan" dan "Yu" berkorelasi dengan topik ajakan (Topik 10), menunjukkan bahwa rekomendasi dan ajakan dari pengguna YouTube dapat mempengaruhi potensi pertumbuhan jumlah pengunjung atau pelanggan baru bagi Lumpia Gang Lombok Semarang. Dengan demikian, melalui analisis ini, dapat dipahami bahwa kata-kata dominan memberikan kontribusi penting untuk memahami berbagai aspek dalam ulasan dan menciptakan gambaran yang lebih lengkap mengenai pengalaman konsumen terhadap Lumpia Gang Lombok Semarang. Dengan memperhatikan aspek-aspek ini dalam pemodelan topik, melalui media sosial YouTube kita dapat lebih mendalam memahami berbagai dimensi yang mempengaruhi opini dan pandangan konsumen terhadap UMKM, memberikan wawasan yang berharga bagi pemilik usaha dan pemirsa yang berpotensi menjadi konsumen. Sehingga, dalam konteks media sosial YouTube sebagai kanal review untuk UMKM, penelitian ini dapat menjadi pertimbangan yang berharga.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa metode Latent Dirichlet Allocation (LDA) dalam Pemodelan Topik dapat berhasil diterapkan pada ulasan media sosial YouTube. Dengan total 524 komentar yang berhasil dikumpulkan, keberhasilan proses pengumpulan data melalui Octoparse 8 menjadi nyata. Tingginya nilai koherensi, terutama dengan mencapai puncak tertinggi sebesar 0,414027, angka ini mengindikasikan bahwa dari keseluruhan sepuluh topik yang dihasilkan, topik ke-10 menunjukkan tingkat konsistensi dan keterkaitan yang paling optimal dengan konten dokumen yang sedang dianalisis. Hal ini menandakan keakuratan dan tingkat interpretabilitas yang baik dalam pemodelan topik menggunakan LDA. Visualisasi intertopic distance map, yang dihasilkan dari pemodelan topik, memberikan gambaran yang jelas tentang kedekatan atau jarak antar topik yang muncul. Dengan kata lain, hasil penelitian ini tidak hanya memberikan kesuksesan dalam analisis topik, tetapi juga memberikan wawasan tentang seberapa terkait atau berbedanya topik-topik yang diidentifikasi dalam ulasan YouTube. Dari hasil pemodelan tersebut, dapat disimpulkan bahwa kanal YouTube dapat dijadikan sumber data yang berharga untuk menilai keinginan masyarakat terhadap suatu produk. Implikasinya, UMKM dapat memanfaatkan hasil analisis ini untuk memahami lebih baik kebutuhan dan keinginan pelanggan, serta merancang strategi pemasaran yang lebih efektif. Dengan memahami pola pikir konsumen melalui platform YouTube, UMKM dapat mengambil keputusan yang lebih cerdas dalam meningkatkan keberhasilan pemasaran produk mereka, termasuk penyesuaian produk, penargetan pelanggan, dan pengembangan strategi pemasaran yang lebih efisien. Secara keseluruhan, penelitian ini tidak hanya berhasil dalam menerapkan metode analisis topik pada data YouTube, tetapi juga memberikan kontribusi signifikan dalam memberikan wawasan yang mendalam dan praktis bagi UMKM untuk mengoptimalkan strategi pemasaran mereka.

UCAPAN TERIMAKASIH

Kami dengan tulus berterima kasih kepada Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, Republik Indonesia yang telah mendanai sebagian



proyek ini melalui Program Kedaireka. Karya ini juga didukung oleh Universitas Dian Nuswantoro (UDINUS) melalui Center of Excellence in Science and Technology, UDINUS dan Asosiasi Pengusaha Oleh-Oleh Jawa Tengah Makanan dan Minuman. dengan hibah dokumen kontrak: Implementasi Supply Chain Management System Pada Umkm Sektor Oleh-Oleh Makanan Dan Minuman, Studi Kasus Di Jawa Tengah No : 4501/F.9.02/UDN- 01/IV/2023

REFERENCES

- [1] M. Sinurat, R. N. Ilham, and W. Cahyadi, "Strategic for Increasing Economic Value Added in the MSME Sector in the North Sumatra Region with the Acceleration Board Method and Initial Public Offering," in Proceedings of the International Conference on Strategic Issues of Economics, Business and, Education (ICoSIEBE 2020), 2021. doi: 10.2991/aebmr.k.210220.019.
- [2] Raden Ariansyah Kamil, "Research on MSMEs in Indonesia : Bibliometric Analysis," *Int J Sci Res Sci Technol*, pp. 171–178, Sep. 2022, doi: 10.32628/IJSRST229527.
- [3] Agni Arifah et al., "Strengthening Business Management and Simple Accounting Understanding for the Rengginang Simanalagi UMKM in Jambuluwuk Village, Ciawi District," *Jurnal Pengabdian Masyarakat Formosa*, vol. 1, no. 5, pp. 557–566, Dec. 2022, doi: 10.55927/jpmf.v1i5.2158.
- [4] S. Tinggi, I. Dakwah, and K. Islam, "KOMUNIKASI DAKWAH MELALUI MEDIA PEMBELAJARAN YOUTUBE Helnafri Ankesa," *TABAYYUN: Jurnal Komunikasi dan Penyiaran Islam*, vol. 3, no. 2, p. 10, 2022, [Online]. Available: <https://ejournal.stmikbogor.ac.id/index.php/tabayyun>
- [5] R. Kurniawan, F. Lestari, A. S. Batubara, M. Z. A. Nazri, K. Rajab, and R. Munir, "Indonesian Lexicon-Based Sentiment Analysis of Online Religious Lectures Review," in 2021 International Congress of Advanced Technology and Engineering (ICOTEN), IEEE, Jul. 2021, pp. 1–5. doi: 10.1109/ICOTEN52080.2021.9493530.
- [6] A. Adak, B. Pradhan, N. Shukla, and A. Alamri, "Unboxing Deep Learning Model of Food Delivery Service Reviews Using Explainable Artificial Intelligence (XAI) Technique," *Foods*, vol. 11, no. 14, p. 2019, Jul. 2022, doi: 10.3390/foods11142019.
- [7] I. K. Susanto, "Analisis Sentimen dan Topic Modelling Pada Pembelajaran Online di Indonesia Melalui Twitter," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 6, no. 2, p. 85, May 2021, doi: 10.31328/jointecs.v6i2.2350.
- [8] Della Maulidiya, "Topic Modelling using Latent Dirichlet Allocation (LDA) to Investigate the Latent Topics of Mathematical Creative Thinking Research in Indonesia," *J. Intell. Comput. Health Inform.*, vol. 3, no. 2, pp. 34–35, 2022, doi: 10.26714/jichi.v3i2.11428.
- [9] A. Andy and U. Andy, "Understanding Communication in an Online Cancer Forum: Content Analysis Study," *JMIR Cancer*, vol. 7, no. 3, p. e29555, Sep. 2021, doi: 10.2196/29555.
- [10] M. Rüdiger, D. Antons, A. M. Joshi, and T.-O. Salge, "Topic modeling revisited: New evidence on algorithm performance and quality metrics," *PLoS One*, vol. 17, no. 4, p. e0266325, Apr. 2022, doi: 10.1371/journal.pone.0266325.
- [11] K. Amaradiena and T. Widarmanti, "SEIKO : Journal of Management & Business LDA-Topic Modeling: Menggunakan Ulasan Pengguna Untuk Meningkatkan User Experience (Studi pada PeduliLindungi)," *SEIKO : Journal of Management & Business*, vol. 6, no. 1, pp. 943–953, 2023, doi: 10.37531/sejaman.v6i1.3802.
- [12] 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, Feb. 2021, doi: 10.29207/resti.v5i1.2587.
- [13] D. Z. T. Kannitha, M. Mustafid, and P. Kartikasari, "PEMODELAN TOPIK PADA KELUHAN PELANGGAN MENGGUNAKAN ALGORITMA LATENT DIRICHLET ALLOCATION DALAM MEDIA SOSIAL TWITTER," *Jurnal Gaussian*, vol. 11, no. 2, pp. 266–277, Aug. 2022, doi: 10.14710/j.gauss.v11i2.35474.
- [14] M. K H, H. Zainuddin, and Y. Wabula, "Twitter Social Media Conversion Topic Trending Analysis Using Latent Dirichlet Allocation Algorithm," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 4, no. 1, pp. 390–399, Dec. 2022, doi: 10.37385/jaets.v4i1.1143.
- [15] M. Habibi, A. Priadana, A. B. Saputra, and P. W. Cahyo, "Topic Modelling of Germas Related Content on Instagram Using Latent Dirichlet Allocation (LDA)," in Proceedings of the International Conference on Health and Medical Sciences (AHMS 2020), Paris, France: Atlantis Press, 2021. doi: 10.2991/ahsr.k.210127.060.
- [16] A. R. Lubis, S. Prayudani, M. Lubis, and O. Nugroho, "Sentiment Analysis on Online Learning During the Covid-19 Pandemic Based on Opinions on Twitter using KNN Method," in 2022 1st International Conference on Information System & Information Technology (ICISIT), IEEE, Jul. 2022, pp. 106–111. doi: 10.1109/ICISIT54091.2022.9872926.
- [17] M. S. Alrajak, I. Ernawati, I. N. Fakultas, and I. Komputer, *ANALISIS SENTIMEN TERHADAP PELAYANAN PT PLN DI JAKARTA PADA TWITTER DENGAN ALGORITMA K-NEAREST NEIGHBOR (K-NN)*. 2020.
- [18] W. Wiranto and Mila Rosyida Uswatunnisa, "Topic Modeling for Support Ticket using Latent Dirichlet Allocation," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 6, pp. 998–1005, Dec. 2022, doi: 10.29207/resti.v6i6.4542.
- [19] A. D. Cahyani and T. Mardiana, "SENTIMENT ANALYSIS OF DIGITAL WALLET SERVICE USERS USING NAÏVE BAYES CLASSIFIER AND PARTICLE SWARM OPTIMIZATION," *Jurnal Riset Informatika*, vol. 2, no. 4, pp. 241–250, Sep. 2020, doi: 10.34288/jri.v2i4.160.
- [20] I. Uglanova and E. Gius, "The Order of Things. A Study on Topic Modelling of Literary Texts," 2020. [Online]. Available: <http://ceur-ws.org>
- [21] S. Mifrah, "Topic Modeling Coherence: A Comparative Study between LDA and NMF Models using COVID'19 Corpus," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5756–5761, Aug. 2020, doi: 10.30534/ijatcse/2020/231942020.