



## *Social Media dan Media Online Analytic Universitas Islam Negeri Mataram dengan Lexicon Based Method dan Latent Dirichlet Allocation*

Ahmad Ashril Rizal <sup>1\*</sup>, Siti Rabi'atul Adawiyah <sup>1</sup>, Anisa Muziyya Rafa <sup>1</sup>

<sup>1</sup> Program Studi Teknologi Informasi, Universitas Islam Negeri Mataram, Indonesia.

\* Korespondensi: [ashril@uinmataram.ac.id](mailto:ashril@uinmataram.ac.id)

**Sitasi:** Rizal, A. A.; Adawiyah, S. R.; Rafa, A. M. (2025). *Social Media dan Media Online Analytic Universitas Islam Negeri Mataram dengan Lexicon Based Method dan Latent Dirichlet Allocation*. JTIM: Jurnal Teknologi Informasi Dan Multimedia, 7(2), 305-325. <https://doi.org/10.35746/jtim.v7i2.620>

Diterima: 24-10-2024

Direvisi: 14-02-2025

Disetujui: 17-03-2025



**Copyright:** © 2025 oleh para penulis. Karya ini dilisensikan di bawah Creative Commons Attribution-ShareAlike 4.0 International License. (<https://creativecommons.org/licenses/by-sa/4.0/>).

**Abstract:** The rapid development of digital technology has revolutionized online and social media analysis, making it an indispensable tool for educational institutions such as Universitas Islam Negeri (UIN) Mataram to comprehend public perception and trending topics. By applying the Latent Dirichlet Allocation (LDA) approach to identify trending topics and the Lexicon-based method to analyze sentiment toward these topics, this research addresses the challenge of identifying key issues and public sentiment. Research data was collected from various news platforms and social media such as Twitter or X, Instagram, and Facebook, using the Google News API and Selenium for web scraping. The collected data was then processed to generate findings relevant to this research. The results of the LDA show that the most frequently discussed topics are related to the achievement of superior accreditation by UIN Mataram, student achievements, and academic activities. Meanwhile, using the lexicon-based method, the sentiment analysis results found that most topics related to UIN Mataram received dominant positive sentiment. Several topics showed positive sentiments, reaching almost 100%, such as topics 2 and 10, which were related to achieving accreditation. Meanwhile, several other issues were dominated by neutral sentiment, indicating that discussions on the topic tended to be informative without triggering significant emotions. Overall, the UIN Mataram institution was viewed positively in the news and media discussions analyzed.

**Keywords:** Media Analytic, UIN Mataram, Latent Dirichlet Allocation (LDA), Sentiment Analysis, Lexicon-Based Method.

**Abstrak:** Pesatnya perkembangan teknologi digital telah merevolusi analisis media sosial dan daring, menjadikannya alat yang sangat diperlukan bagi lembaga pendidikan seperti Universitas Islam Negeri (UIN) Mataram untuk memahami persepsi publik dan topik yang sedang tren. Dengan menerapkan pendekatan Latent Dirichlet Allocation (LDA) untuk mengidentifikasi topik yang sedang tren dan Lexicon-Based Method untuk menganalisis sentimen terhadap topik tersebut, penelitian ini menjawab tantangan untuk mengidentifikasi isu-isu utama dan sentimen publik. Data penelitian dikumpulkan dari berbagai platform berita dan media sosial seperti Twitter atau X, Instagram, dan Facebook, menggunakan Google News API dan Selenium untuk web scraping. Data yang terkumpul kemudian diolah untuk menghasilkan temuan yang relevan dengan penelitian ini. Hasil LDA menunjukkan bahwa topik yang paling sering dibahas terkait dengan pencapaian akreditasi unggul oleh UIN Mataram, prestasi mahasiswa, dan kegiatan akademik. Sementara itu, dengan menggunakan dan Lexicon-Based Method, hasil analisis sentimen menemukan bahwa

sebagian besar topik yang terkait dengan UIN Mataram menerima sentimen positif yang dominan. Beberapa topik menunjukkan sentimen positif, mencapai hampir 100%, seperti topik 2 dan 10 yang terkait dengan perolehan akreditasi. Sementara itu, beberapa isu lainnya didominasi sentimen netral, yang menunjukkan bahwa pembahasan pada topik tersebut cenderung informatif tanpa memicu emosi yang berarti. Secara keseluruhan, institusi UIN Mataram dinilai positif dalam berita dan diskusi media yang dianalisis.

**Kata Kunci:** Media Analytic, UIN Mataram, Latent Dirichlet Allocation (LDA), Sentiment Analysis, Lexicon-Based Method.

## 1. Pendahuluan

Media analitik telah menjadi salah satu alat yang sangat berpengaruh dalam memahami tren topik dan sentimen yang berkembang di dunia digital, terutama melalui platform media sosial dan media online [1]. Promosi yang dilakukan oleh berbagai pihak untuk mencapai tujuan pemasaran kini mulai bergeser dari konvensional ke digital, bahkan media promosi pun dilakukan dengan menggunakan konsep *augmented reality* [2]. Bahkan pada saat pandemi covid-19, media digital menjadi penyebaran informasi yang paling efektif untuk sosialisasi dan koordinasi penanganan covid-19 pada saat itu [3]. Media analisis pada sosial media juga kini lebih mudah diterapkan untuk diambil sebagai dataset melalui website [4]. Social media analytic framework (SMAF) adalah salah satu metode media analitik yang diintegrasikan dengan regresi fuzzy. SMAF digunakan untuk mengekstrak dan menganalisis data desain afektif dari platform media sosial seperti Twitter, Facebook, dan Instagram [5].

Pada institusi pendidikan, kehumasan merupakan salah satu unit yang sangat penting dan berperan sebagai penyebaran informasi. Selama ini, tim kehumasan universitas telah bekerja keras dalam menyosialisasikan berbagai program dan pencapaian akademik melalui media sosial dan media online. Namun, belum pernah dilakukan analisis yang sistematis terkait dampak komunikasi kehumasan tersebut terhadap persepsi publik. Tanpa analisis berbasis data, universitas belum dapat mengetahui sejauh mana efektivitas strategi komunikasi yang diterapkan, terutama dalam membangun citra positif di kalangan masyarakat. Oleh karena itu, penelitian ini menjadi sangat penting dalam membantu universitas memahami bagaimana sosialisasi program akademik dan pencapaian institusi diterima oleh publik, serta memberikan rekomendasi berbasis data untuk meningkatkan strategi komunikasi yang lebih efektif. Di era digital khususnya, persepsi masyarakat terhadap suatu institusi sangat dipengaruhi oleh media sosial dan pemberitaan online, yang dapat membentuk citra positif maupun negatif. Tanpa metode analitik yang tepat, universitas dapat kesulitan dalam mengidentifikasi isu-isu yang berkembang dan merancang strategi komunikasi yang efektif.

Analisis media ini menjadi aspek penting karena memungkinkan organisasi, termasuk institusi pendidikan seperti Universitas Islam Negeri (UIN) Mataram, untuk lebih memahami persepsi publik serta isu-isu utama yang sedang ramai dibicarakan. Dengan berkembangnya teknologi informasi, berbagai sektor termasuk pendidikan mulai mengadopsi metode *media analytic* untuk mengukur sentimen publik dan menganalisis topik yang berkembang, yang kemudian digunakan untuk mengambil keputusan strategis [6]. Analisis media digital memungkinkan identifikasi tren topik yang banyak dibicarakan oleh masyarakat. Dalam konteks UIN Mataram, pemahaman yang mendalam mengenai isu-isu yang sering diperbincangkan dan bagaimana persepsi publik terhadap universitas merupakan langkah yang sangat penting untuk menjaga reputasi lembaga pendidikan ini. Selain itu, melalui pendekatan ini, UIN Mataram dapat mengembangkan strategi komunikasi yang lebih baik untuk merespons isu-isu yang sedang berkembang di kalangan masyarakat, terutama di ranah online.

Analisis media digital memungkinkan identifikasi tren topik yang banyak dibicarakan oleh masyarakat. Dalam konteks UIN Mataram, pemahaman yang mendalam mengenai isu-isu yang sering diperbincangkan dan bagaimana persepsi publik terhadap universitas merupakan langkah yang sangat penting untuk menjaga reputasi lembaga pendidikan ini. Selain itu, melalui pendekatan ini, UIN Mataram dapat mengembangkan strategi komunikasi yang lebih baik untuk merespons isu-isu yang sedang berkembang di kalangan masyarakat, terutama di ranah online. Penelitian ini melakukan pendekatan berbasis *Latent Dirichlet Allocation* untuk mengidentifikasi topik yang sedang tren dan *Lexicon-Based Sentiment Analysis* untuk mengukur sentimen publik terhadap isu-isu tersebut. Dengan hasil yang diperoleh, universitas dapat mengambil langkah strategis dalam meningkatkan citra akademik dan merespons opini publik secara lebih terukur.

Salah satu metode yang banyak digunakan dalam analisis topik adalah Latent Dirichlet Allocation (LDA), sebuah algoritma yang berfungsi untuk menemukan pola-pola tersembunyi atau topik dari kumpulan teks yang besar [7]. LDA memungkinkan para peneliti untuk melihat tren topik yang muncul tanpa memerlukan label manual pada data teks [8] [9]. Dalam penelitian ini, LDA digunakan untuk mengidentifikasi topik-topik yang sering muncul dalam diskusi di media *online* dan media sosial yang terkait dengan UIN Mataram. Metode ini berguna untuk memisahkan tema-tema utama yang mungkin menjadi perhatian publik dan memberikan wawasan yang lebih mendalam mengenai topik yang dominan dalam percakapan daring. Selain analisis topik, pemahaman tentang sentimen publik juga penting. Sentimen publik menggambarkan emosi dan opini yang berkembang terhadap suatu topik. Analisis sentiment ini bisa dilakukan menggunakan algoritma NLP (*Natural Language Processing*) [6], [10]. Salah satu teknik atau metode NLP yang sering digunakan dalam analisis sentimen yaitu *lexicon-based* atau berbasis leksikon. Metode ini bekerja dengan mencocokkan kata-kata dalam teks dengan daftar kata-kata yang sudah diklasifikasikan sebagai positif, negatif, atau netral. Dengan menggunakan kamus kata positif dan negatif, metode ini dapat menentukan apakah opini yang disampaikan dalam teks cenderung positif, negatif, atau netral [11]. Pendekatan ini sederhana namun efektif dalam menganalisis teks yang tidak berstruktur, seperti komentar atau ulasan di media sosial [12].

Sejumlah penelitian terdahulu mendukung keefektifan penggunaan LDA dan analisis sentimen berbasis leksikon dalam analisis teks media sosial dan media online. Misalnya, penelitian oleh [13] menggunakan metode LDA dan analisis sentimen berbasis leksikon untuk memahami opini publik terhadap topik-topik yang sedang tren di Provinsi Nusa Tenggara Barat. Penelitian ini menunjukkan bahwa kombinasi antara LDA dan metode *lexicon-based* dapat memberikan pemahaman yang mendalam tentang topik dan emosi publik terhadap isu-isu lokal yang relevan. Penelitian lain oleh [14] menegaskan bahwa LDA sangat efektif dalam mengekstraksi topik dari laporan teks, sementara analisis sentimen berbasis leksikon memberikan wawasan yang jelas tentang bagaimana masyarakat bereaksi terhadap berbagai isu yang muncul. Penelitian ini sendiri bertujuan untuk menerapkan pendekatan LDA dan metode *lexicon-based* pada data media *online* dan media sosial yang terkait dengan UIN Mataram. Penelitian ini bertujuan untuk memberikan wawasan mengenai topik-topik utama yang dibicarakan terkait UIN Mataram, serta bagaimana persepsi publik terhadap topik-topik tersebut, baik positif, negatif, atau netral. Melalui analisis ini, UIN Mataram diharapkan dapat mengidentifikasi topik yang perlu diperhatikan dan mengukur sentimen publik untuk meningkatkan manajemen komunikasi dan hubungan masyarakat secara keseluruhan.

Kombinasi metode LDA dan analisis sentimen berbasis leksikon yang jarang digunakan secara bersamaan dalam kajian media digital untuk institusi pendidikan. Studi ini tidak hanya mengidentifikasi tren topik di media online dan sosial, tetapi juga mengukur sentimen masyarakat terhadap topik-topik tersebut dan memberikan pemahaman yang lebih mendalam. Selain itu, penelitian ini menerapkan pendekatan

multi-platform dengan mengumpulkan data dari Google News, Twitter/X, Instagram, dan Facebook, menggunakan metode web scraping dengan Selenium dan Google News API. Hal ini menjadikannya lebih komprehensif dibandingkan penelitian lain yang hanya berfokus pada satu platform.

Penggunaan metode LDA dan analisis sentimen berbasis leksikon tidak hanya membantu dalam menangkap tren dan topik yang sedang berkembang, tetapi juga memberikan wawasan lebih lanjut tentang bagaimana opini publik terhadap topik tersebut terbentuk. Misalnya [13] berhasil menggunakan metode ini untuk menganalisis *trending topics* di media sosial terkait isu-isu di Provinsi Nusa Tenggara Barat. Hasil dari analisis tersebut mampu memberikan wawasan yang lebih mendalam kepada pemerintah daerah mengenai opini publik terhadap isu-isu lokal yang mereka hadapi. Dalam penelitian ini, pendekatan yang sama digunakan untuk UIN Mataram dengan harapan dapat membantu universitas dalam memahami lebih dalam mengenai persepsi publik terhadapnya.

## 2. Metode Penelitian

Penelitian ini mencoba melakukan analisis media digital yang berfokus pada Universitas Islam Negeri (UIN) Mataram. Tujuan dari penelitian ini adalah untuk memahami persepsi publik terhadap UIN Mataram melalui analisis topik yang sedang tren dan sentimen yang berkembang di berbagai *platform* media sosial serta media *online*. Analisis ini sangat penting untuk membantu UIN Mataram dalam mengelola citra publik, meningkatkan komunikasi strategis, dan memahami isu-isu utama yang dibicarakan oleh masyarakat. Tujuan utama dari analisis ini adalah untuk mengetahui sentimen publik terkait UIN Mataram di media sosial, mengidentifikasi tren atau isu-isu yang sedang populer terkait UIN Mataram, menilai *engagement* masyarakat terkait postingan dan berita tentang UIN Mataram, dan mengukur persepsi publik tentang layanan atau acara yang diselenggarakan oleh UIN Mataram.

Setelah tujuan ditetapkan, maka ditentukan metrik yang sesuai untuk diukur dari berbagai *platform*. Metrik ini mencakup Twitter/ X dalam menganalisis teks yang diperoleh dari *tweet* dengan kata kunci "UIN Mataram" atau hashtag #UINMataram. Menganalisis teks yang diperoleh dari *caption* yang mengandung kata kunci "UIN Mataram" atau *hashtag* #UINMataram di Instagram. Menganalisis postingan yang mengandung kata kunci "UIN Mataram" atau *hashtag* #UINMataram di facebook. Dan mengambil artikel dari media *online* yang mengandung kata kunci "UIN Mataram" melalui Google News API.

### Collecting Data

Data dikumpulkan dari dua sumber utama, yaitu media online dan media sosial. Proses pengumpulan data mencakup:

1. Media Online dengan Google News API:  
Artikel-artikel yang berkaitan dengan UIN Mataram diambil melalui *Google News API*. Dalam proses ini, kata kunci seperti "Universitas Islam Negeri Mataram" digunakan untuk mengidentifikasi artikel yang relevan. Artikel yang diperoleh disimpan dalam format JSON atau CSV untuk dianalisis lebih lanjut.
2. Twitter/X:  
Data dikumpulkan menggunakan Selenium untuk mengotomatisasi pencarian *tweet* yang mengandung kata kunci "UIN Mataram". Hasil scraping kemudian disimpan dalam format CSV atau JSON untuk dianalisis lebih lanjut.
3. Instagram:  
Postingan di Instagram yang mengandung hashtag #UINMataram diambil dengan menggunakan Selenium. Hasil dari scraping ini juga disimpan dalam format CSV atau JSON untuk dianalisis lebih lanjut.

#### 4. Facebook:

Postingan di Facebook yang berkaitan dengan UIN Mataram juga dikumpulkan menggunakan Selenium. Proses scraping pada platform ini dilakukan dengan mencari kata kunci yang relevan dan menyimpan hasilnya untuk dianalisis.

#### **Trending Topic dengan LDA**

Langkah berikutnya adalah menggunakan metode Latent Dirichlet Allocation (LDA) untuk menemukan topik-topik yang sedang tren. LDA adalah algoritma untuk menemukan pola tersembunyi atau topik yang sering muncul dalam dokumen teks. Setiap dokumen dianggap sebagai campuran dari beberapa topik, dan setiap topik dianggap sebagai distribusi dari kata-kata. Proses LDA dalam penelitian ini melibatkan beberapa langkah [15]:

1. Preprocessing Teks: Mengubah teks menjadi huruf kecil, menghapus tanda baca, menghilangkan stopwords, dan melakukan stemming untuk mengambil akar kata.
2. TF-IDF: Mengubah teks yang sudah diproses menjadi representasi numerik berdasarkan frekuensi kata (Term Frequency - Inverse Document Frequency).
3. Latent Dirichlet Allocation: Setelah teks diubah menjadi representasi numerik, model LDA digunakan untuk mengidentifikasi 10 topik yang paling sering muncul dalam diskusi terkait UIN Mataram.
4. Visualisasi Hasil: Topik yang dihasilkan oleh LDA ditampilkan dalam bentuk daftar kata-kata kunci yang paling mewakili topik tersebut.

#### **Sentiment Analysis dengan Lexicon-Based Method**

Setelah menemukan topik yang sedang tren, langkah selanjutnya adalah melakukan analisis sentimen terhadap topik-topik tersebut menggunakan metode berbasis leksikon (Lexicon-Based Method). Metode ini menggunakan kamus kata positif dan negatif untuk menentukan apakah suatu teks memiliki sentimen positif, negatif, atau netral. Proses analisis sentimen dilakukan dengan langkah-langkah berikut:

1. Preprocessing Teks: Sama seperti pada LDA, teks terlebih dahulu diproses untuk menghilangkan karakter yang tidak relevan, menghapus stopwords, dan melakukan stemming.
2. Pencocokan Kata dengan Kamus Sentimen: Setelah teks diproses, setiap kata dalam teks diperiksa apakah ada dalam kamus kata positif atau negatif. Jika kata tersebut ada dalam kamus, maka nilainya ditambahkan ke skor sentimen total.
3. Klasifikasi Sentimen: Setelah seluruh kata dalam teks dianalisis, sentimen dari setiap dokumen diklasifikasikan sebagai positif, negatif, atau netral berdasarkan skor totalnya [10].

### **3. Hasil**

#### *3.1. Collecting Data*

Hasil dataset disimpan dalam format csv (*comma separated values*). Dataset diambil pada periode 1 Mei 2024 – 30 September 2024.

##### *3.1.1. Media Online*

Tabel 1 berikut menyajikan 10 *dataset* terbaru yang diambil dari *googlenews*.

**Tabel 1.** *Dataset dari Google News*

Title	Media
Rektor UIN Mataram Apresiasi Peningkatan Indeks Kepuasan Jemaah Haji Metro NTB Indonesia - Metro NTB	
FTK UIN Mataram Gelar Upgrading Mahasiswa - Metro NTB	Metro NTB

Title	Media
Alumninya Paling Banyak Diterima UIN Mataram, MAN 1 Mataram Terima Penghargaan Rektor - Lombok Post	Lombok Post
SPI UIN Mataram Gelar Workshop Manajemen Risiko dan Penyusunan Peta Risiko - Metro NTB	Metro NTB
FTK UIN Mataram Gelar Diskusi Lintas Disipliner bersama Dosen UPSI Malaysia - Metro NTB	Metro NTB
PGMI UIN Mataram dan UPSI Malaysia, Asah Kreativitas Calon Guru SD/MI dalam Pembelajaran Seni - Metro NTB	Metro NTB
Masyarakat kecanduan berolahraga di RTH Pagutan Mataram	ANTARA News
Mahasiswa, Agen Perubahan dalam Pilkada 2024	<a href="https://rri.co.id">RRI.co.id</a>
Mahasiswa Program Studi Bimbingan dan Konseling Islam Melaksanakan Praktik Kerja Lapangan (PKL) Reguler FDIK UIN Mataram di Bale Mediasi Kota Mataram - Kompasiana.com	<a href="https://kompasiana.com">Kompasiana.co</a>
Akademisi UIN Mataram ungkap kekuatan media sosial pengaruhi opini publik pilkada	ANTARA News

### 3.1.2. Twitter

Dataset yang diperoleh dari twitter dengan menggunakan selenium ditunjukkan pada Tabel 2. Pada Tabel 2 ditampilkan 10 *dataset* yang diperoleh dari twitter atau X.

**Tabel 2.** Dataset dari Twitter atau X

akun	tanggal	tweet
ATCS @ATCS_DISHUB _DIY . Sep 13	DIY 9/15/2024	Pantauan arus lalu lintas Tanggal 13 September 2024 Pukul 07.00 WIB terpantau dari : 1. Sp Janti dari Barat (Kiri atas) Lancar 2. Sp UIN dari Timur (Kanan Atas) Ramai Lancar 3. Sp Selokan Mataram dari Selatan(Kiri Bawah) Ramai Lancar 4. Sp Wojo dari Timur (Kanan Bawah) Lancar
arina.id @arinadotid . Sep 3	9/15/2024	3 PTKIN yakni UIN Mataram, UIN KHAS Jember dan UIN Gus Dur Pekalongan meraih akreditasi unggul dari BAN PT. Direktur Diktis Ahmad Zainul Hamdi
arina.id @arinadotid . Aug 15	9/15/2024	Pemkab Lombok Barat NTB menghibahkan tanah untuk sarana pendidikan UIN Mataram. Tanah seluas 3,1 hektar ini memakan jarak tempuh 10 menit berkendara dari kampus UIN.
Amihtndsla @athumilah11644 . Jul 17	9/15/2024	MasyaAllah udah sampai di tahap ini #yudisiumuinmataram #uinmataram
HADY @Hadykix22 . Jul 3	9/15/2024	Udh tau dari dulu min salam anak uin #UINmataram
Pascasarjana UIN Mataram @PascasarjanaUin . Jun 11	9/15/2024	Yuk cari tau Program Studi apa saja yang ada di Pascasarjana UIN Mataram  Info Seputar PMB dan Pendaftaran dapat diakses melalui Scan QR atau dengan mengunjungi laman

akun	tanggal	tweet
		<a href="https://pmb.pascasarjana.uinmataram.ac.id">https://pmb.pascasarjana.uinmataram.ac.id</a>
		#pmb #pmbuinmataram #uinmataram #pascasarjana #uin #infopmb #beasiswa #bib
Suhai(ri) Ahmad #UsutTuntasKanj urukan @suhairi4hmad . Apr 7	D. 9/15/2024 	Ngomong-ngomong soal arkeologi makam-makan kuno, aku ingat buku seorang teman yang sekaligus dosen di UIN Jogja. Ia fokus meneliti makam-makan tua sejak era Mataram Islam.
agus santhosa @agus_santhosa . Feb 24	9/15/2024	Negara harus bersahabat dengan alam. Mahasiswa Prodi KPI UIN Mataram go green di G. Aur, Desa Sedau, Narmada, Lombok Barat
Merapi Uncover @merapi_uncover . Jan 15	9/15/2024	Siapa tau ada yg menemukan lanyard ungu ada tulisannya "CE Engagement" dari area Alkid, Jl. Mataram, Std. Kridosono, SMP N 5 YK, Jl. Kusbini (Balai Yasa YK), Kampus Akprind, SD Muh. Sapen, UIN Sunan Kalijaga, Mall Lippo Yk.
		Bsa menghubungi 089674082900 Kayla.   @dinodtans
Monly @marsieles_monl y . Dec 24, 2023	9/15/2024	Kerja sama KPU dan UIN Mataram untuk memperkuat KPU dalam memberikan layanan kepada para pemilih untuk jangkauan yang semakin luas #PemiluJujur #PemiluIndonesia #Pemilu2024 #Pemilu Pagii Subuh Tidur Wa Aja #JagaNetralitas
Regarcommerce @Erwinbinjai . Nov 16, 2023	9/15/2024	panorama pemikiran islam ulama maroko - Repository UIN Mataram <a href="http://repository.uinmataram.ac.id/2345/1/PANORAMA%20PEMIKIRAN%20ISLAM%20ULAMA%20MAROKO.pdf">http://repository.uinmataram.ac.id/2345/1/PANORAMA%20PEMIKIRAN%20ISLAM%20ULAMA%20MAROKO.pdf</a> ?

### 3.1.3. Instagram

Hasil crawling dataset pada sosial media instagram dengan keyword “UIN Mataram” ditunjukkan pada Tabel 3. Pada Tabel 3 ditampilkan 10 dataset terbaru daru hasil crawling data dengan menggunakan selenium.

**Tabel 3.** Dataset dari Instagram

akun	tanggal	caption
masuk_p tkin	2024-09-06	Pengen masuk PTKIN impian kamu? Tapi bingung mulai belajar dari mana? dan belum punya info seputar ujian masuk PTKIN? 😊
		Buruan follow @masuk_ptkin karena bakal ada latihan soal, tips & trik belajar, dan info lengkap soal ujian masuk PTKIN setiap harinya!
		Belajar jadi lebih mudah dan menyenangkan, plus kamu tetap update informasi terbaru seputar ujian PTKIN, hanya di @masuk_ptkin!
		Jangan sampai ketinggalan! Yuk, follow sekarang supaya kamu lebih siap

akun	tanggal	caption
		menghadapi ujian dan meraih PTKIN impianmu! ✨
		@masuk_ptkin TEMAN AMBISMU!!!! 🤖
		#umptkin #umptkin2025 #infoumptkin2025 #infoumptkin #tryoutonlineumptkin #tryoutpremium #tryoutonline #uinjakarta #uinbandung #uinsemarang #uinjogja #uinsumatera #uinradenfatah #uinsaizu #uinradenmassaid #uinmakassar #uinmalikimalang #uinsayyidalirahmatullah #uinmataram #iainparepare #uinsalatiga
		Coba, pertanyaan yang diatas jawabannya apa ya??? Jawaban dan pembahasannya bakal di spill besok, jadi stay tune 😊
dea.lutfia naa	2024-09-06	"Gabung bersama kami di UIN Mataram, Khususnya program studi komunikasi dan penyiaran islam"
		#uinmataram #kpi #kecebadai #keren #pateen #menyala #indonesia
kepegaw aian.uin mataram	2024-09-06	Assalamualaikum Warahmatullahi Wabarakatuh  Alhamdulillah, Tahniah Universitas Islam Negeri Mataram telah meraih predikat akreditasi UNGGUL berdasarkan keputusan Badan Akreditasi Nasional Perguruan Tinggi pada 3 September 2024.  Semoga makin berjaya dan sukses selalu. aamiin Wassalamu'alaikum Warahmatullahi Wabarakatuh
		<b>Organisasi UIN dan Kepegawaian Mataram</b>
		<b>Kami Ada untuk Anda</b>
		#ekinerja #kepegawaianuinmataram #uinma #perguruan tinggi #uin #uinma #uinmataram #ikhlasberamal #kemenagri
societion _a23	2024-09-06	#mahasiswa #uinmataram #fypシ #beranda #uinma
masuk_p tkin	2024-09-06	Pengen masuk PTKIN impian kamu? Tapi bingung mulai belajar dari mana? dan belum punya info seputar ujian masuk PTKIN? 🤔  Buruan follow @masuk_ptkin karena bakal ada latihan soal, tips & trik belajar, dan info lengkap soal ujian masuk PTKIN setiap harinya!  Belajar jadi lebih mudah dan menyenangkan, plus kamu tetap update informasi terbaru seputar ujian PTKIN, hanya di @masuk_ptkin!  Jangan sampai ketinggalan! Yuk, follow sekarang supaya kamu lebih siap menghadapi ujian dan meraih PTKIN impianmu! ✨
		@masuk_ptkin TEMAN AMBISMU!!!! 🤖
		#umptkin #umptkin2025 #infoumptkin2025 #infoumptkin #tryoutonlineumptkin #tryoutpremium #tryoutonline #uinjakarta #uinbandung #uinsemarang #uinjogja #uinsumatera #uinradenfatah #uinsaizu #uinradenmassaid #uinmakassar #uinmalikimalang #uinsayyidalirahmatullah #uinmataram #iainparepare #uinsalatiga



[illegible]

## 3.1.4. Facebook

Dataset dari facebook juga diambil dengan menggunakan selenium. Hasil crawling dataset ditunjukkan dalam Tabel 4. Tabel 4 berisi 10 data sampel hasil crawling data facebook dengan keyword “UIN Mataram”.

Tabel 4. Dataset dari Facebook

akun	tanggal	status
Fathurrahman Jr.	2024-09-06	Barakallah UIN Mataram. #Sukses #berkah #unggul
Sentra Paramita	2024-09-06	Hai Sobat Sosial... Hari ini (05/09), Sentra Paramita kedatangan tamu dari Universitas Islam Negeri (UIN) Mataram nih, tepatnya dari jurusan Bimbingan Konseling Islam. Diterima langsung oleh Kepala Sentra Paramita Mataram didampingi oleh Kasubbag TU Sentra Paramita Mataram, seluruh mahasiswa yang didampingi oleh dosen pembimbingnya, Bapak Syamsul Hadi, mahasiswa UIN tersebut akan melakukan Praktik Kerja Lapangan (PKL) selama 45 hari di Sentra Paramita Mataram. Hal mendas... Lihat selengkapnya
Mu Din	2024-09-06	Sedang mengunjungi keponakan yang Kuliah di UIN Mataram(Lebih tepatnya Di Jempong) berdekatan dengan Kampus UIN Mataram.
Maimun Zubair	2024-09-06	Terima kasih kepada mitra kerja UIN Mataram atas ucapannya. Capaian Unggul untuk Akreditasi Instutusi tidak lepas dari dukungan mitra kerja dan stakeholders.[]
M Sobry Sutikno	2024-09-06	Alhamdulillah uin mataram dapat akreditasi unggul. Jangan ragu bapak ibu untuk kuliahkan putra putri atau sanak saudaranya di uin mataram...
Riduan Mas'ud	2024-09-06	Ke pantai Senggigi naik perahu, Angin sepoi membawa damai, UIN Mataram unggul tak diragukan lagi, Satu-satunya di NTB yang terus memulai.
Putri Nabila Aulia	2024-09-06	Sebangga itu pernah kuliah di UIN Mataram akreditasi Unggul menyala UIN ku
I Wayan Wirata	2024-09-06	Selamat dan sukses atas pencapaian akreditasi ungu UIN Mataram,
Asriadi Azhar	2024-09-06	Trimakasih kepada keluarga besar MI Nurul Karim Kebon Ayu Lombok Barat atas kerjasamanya dalam mensukseskan kegiatan Pengabdian Masyarakat Prodi PGMI UIN Mataram. Smoga senantiasa dalam keberkahan. Aamiin
Hamdan i Khaerul Fikri	2024-09-06	Torehan sejarah yang membanggakan bagi seluruh Civitas Akademika UIN Mataram Khususnya kepada seluruh Jajaran Pimpinan UIN Mataram sehingga UIN Mataram menjadi satu-satunya Universitas yang meraih Predikat Unggul di NTB. Berkah selalu UIN Mataram dalam membangun Generasi Anak bangsa yg membanggakan – di Kampus II UIN ( Universitas Islam Negeri ) MATARAM.
Miski AhLa Jinan	2024-09-06	Apresiasi untuk PascaSarjana UIN MATARAM yg Baru saja Mendapatkan Predikat Akreditasi UngguL (A), Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi, Pada 3 September 2024. #Yuk_S1_S2_S3_di_UIN_MATARAM #Me_Start_S2_IQT_PascaSarjana

### 3.2. *Preprocessing*

#### 3.2.1. Normalisasi

Langkah *preprocessing* yang pertama adalah membuat kumpulan *text* pada *dataset* menjadi huruf kecil.

#### 3.2.2. Stop Word

Langkah *preprocessing* berikutnya adalah menghapus *character* yang *non-alphabet* karena tidak memiliki makna.

#### 3.2.3. Stemming

*Stemming* adalah langkah berikutnya dari *preprocessing* yang berperan untuk mengubah kata imbuhan menjadi kata dasar. Pada penelitian ini *library* yang digunakan adalah Sastrawi. Sastrawi adalah sebuah *library Python* yang digunakan untuk melakukan *stemming* pada teks berbahasa Indonesia. *Stemming* adalah proses mengubah kata-kata terinfleksi atau kata turunan (seperti awalan, akhiran, atau imbuhan) menjadi bentuk dasar atau kata dasar (*root word*). Proses ini sangat penting dalam pengolahan bahasa alami (*Natural Language Processing/NLP*), terutama untuk tugas seperti analisis sentimen, text mining, dan klasifikasi teks.

Fitur utama Sastrawi untuk *stemming* Kata Bahasa Indonesia adalah menghapus berbagai imbuhan dalam bahasa Indonesia seperti awalan (me-, di-, ber-, ke-), sisipan, dan akhiran (-kan, -i, -an) untuk mendapatkan bentuk dasar kata.

#### 3.2.4. Tokenization

*Tokenization* adalah proses memecah teks menjadi unit-unit kecil seperti kata, kalimat, atau bahkan karakter. *Tokenization* merupakan langkah dasar dalam pengolahan bahasa alami (*Natural Language Processing*). Langkah dalam potongan kode untuk melakukan *tokenization* di Python menggunakan library *nlTK* (*Natural Language Toolkit*).

## 4. Pembahasan

Hasil *trending topic* ditentukan dengan menggunakan metode LDA (*Latent Dirichlet Allocation*).

### 4.1. *Trending Topic*

#### 4.1.1. *Text Frequency*

Menghitung frekuensi kemunculan dalam suatu dataset text dapat dihitung dengan menggunakan library *CountVectorizer* dari *scikit-learn*.

Hasil dari menghitung frekuensi text yang muncul ditunjukkan pada Gambar 1.

```

Mounting Google Drive...
Mounted at /content/drive
Loading file...
Extracting bigrams and trigrams...

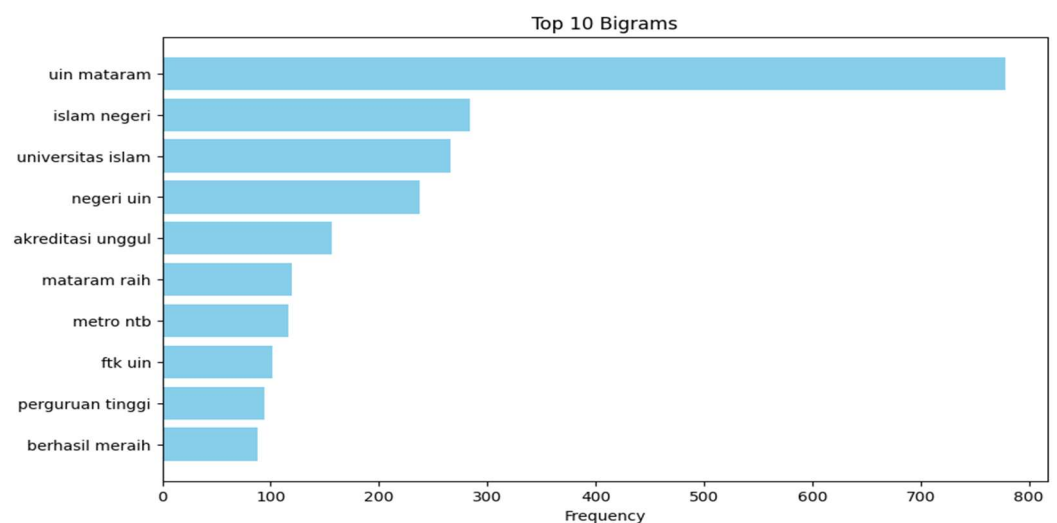
Top Bigrams:
uin mataram: 778
islam negeri: 284
universitas islam: 266
negeri uin: 238
akreditasi unggul: 156
mataram raih: 120
metro ntb: 116
ftk uin: 102
perguruan tinggi: 94
berhasil meraih: 88

Top Trigrams:
universitas islam negeri: 266
islam negeri uin: 238
negeri uin mataram: 238
uin mataram raih: 120
ftk uin mataram: 102
uin mataram berhasil: 82
berhasil meraih predikat: 80
uin mataram gelar: 80
mataram berhasil meraih: 72
mataram raih akreditasi: 64

```

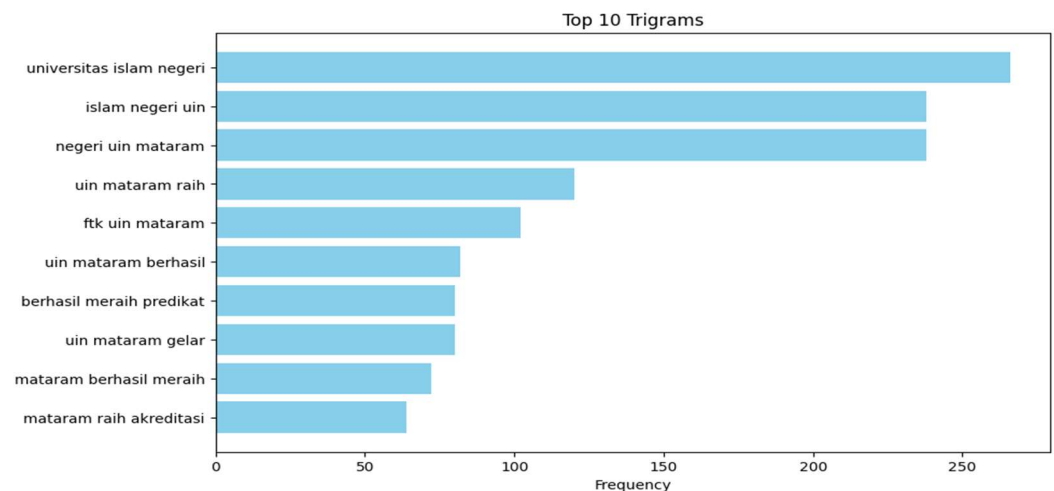
**Gambar 1.** Hasil Top Bigram dan Trigram

Program yang dibuat akan menjalankan fungsi untuk mengekstrak bigram (dua kata berurutan) dan trigram (tiga kata berurutan) dari teks yang ada dalam file CSV. Dari hasil tersebut, daftar bigram (dua kata yang muncul secara berurutan) yang paling sering muncul dalam teks adalah "uin mataram" muncul sebanyak 778 kali dalam teks, "islam negeri" muncul sebanyak 284 kali. Hasil selengkapnya ditunjukkan dalam grafik pada Gambar 2 yang menampilkan frekuensi bi-gram pada *dataset*.



**Gambar 2.** Grafik Frekuensi Bigram

Bigram membantu mengidentifikasi pasangan kata yang sering digunakan bersama dalam teks yang dianalisis. Sementara itu, trigram (tiga kata yang muncul secara berurutan) yang paling sering muncul dalam teks ditunjukkan dalam Gambar 3. Frasa "universitas islam negeri" muncul sebanyak 266 kali. Trigram "islam negeri uin" muncul sebanyak 238 kali. Detail dari kemunculan Trigram dapat dilihat pada Gambar 3.



**Gambar 3.** Grafik Frekuensi Trigram

Bigram dan trigram yang muncul berkaitan erat dengan topik yang ada dalam teks. Berdasarkan bigram dan trigram ini, tampak jelas bahwa teks yang dianalisis banyak berbicara tentang UIN Mataram, akreditasi unggul, dan beberapa hal terkait universitas dan pendidikan tinggi di Indonesia. UIN Mataram dan universitas islam negeri mendominasi, menandakan bahwa ini adalah topik utama dalam korpus teks. Bigram dan trigram ini bisa digunakan untuk memahami tema utama atau fokus dari teks, misalnya berita atau artikel yang banyak menyoroti pencapaian akreditasi dan nama institusi pendidikan seperti UIN Mataram.

#### 4.1.2. Latent Dirichlet Allocation

*Latent Dirichlet Allocation* (LDA) adalah salah satu metode topik *modeling* yang digunakan untuk menemukan tema atau topik tersembunyi di dalam sekumpulan dokumen teks. Dalam konteks mencari trending topik, LDA sangat berguna karena dapat mengidentifikasi berbagai tema atau isu yang sedang dibahas secara dominan dalam sejumlah besar teks.

##### a) Preprocessing

LDA bekerja dengan cara mengelompokkan kata-kata yang sering muncul bersama-sama ke dalam beberapa topik, kemudian memberikan bobot untuk setiap dokumen berdasarkan seberapa besar kemunculan topik tersebut dalam dokumen tersebut. Dengan demikian, LDA dapat membantu mengidentifikasi topik atau tren yang muncul secara signifikan dalam suatu periode waktu tertentu.

LDA membantu menemukan isu-isu utama yang sedang dibahas secara dominan dalam media sosial atau berita online. Topik yang relevan dapat ditemukan secara otomatis dari kumpulan data tanpa perlu mengidentifikasi atau menetapkan topik sebelumnya. LDA juga dapat digunakan untuk memantau perubahan dalam tren topik dari waktu ke waktu. Langkah awal dari LDA adalah membersihkan dataset agar dapat diproses.

Fitur *stopwords* diunduh dengan NLTK (`nltk.download('stopwords')`). Dataset yang berisi artikel berita diambil dari file CSV yang disimpan di Google Drive. Dataset ini berisi kolom title yang berisi judul-judul berita yang akan dianalisis.

Pada proses *lowercasing*, semua huruf dalam teks diubah menjadi huruf kecil agar tidak ada perbedaan antara kata-kata dengan huruf kapital dan non-kapital. Kemudian semua tanda baca dihilangkan untuk mencegah elemen yang tidak diperlukan dalam pemrosesan teks. Proses tokenisasi dan *stopwords* akan memecah kata-kata (token) dan menghapus kata-kata yang dianggap tidak penting.

## b) TF-IDF

Proses TF-IDF dilakukan dengan menggunakan `TfidfVectorizer` untuk mengubah teks menjadi representasi numerik berdasarkan frekuensi kata dengan mempertimbangkan seberapa umum kata tersebut dalam seluruh dokumen. Nilai `max_df=0.95` artinya proses akan mengabaikan kata-kata yang muncul di lebih dari 95% dokumen (kata terlalu umum). Nilai `min_df=2` artinya proses akan mengabaikan kata-kata yang muncul kurang dari dua dokumen (kata terlalu jarang).

TF-IDF membantu memberikan bobot pada kata-kata yang tidak hanya sering muncul dalam satu dokumen, tetapi juga memperhitungkan apakah kata tersebut sering muncul di seluruh dokumen. Semakin umum kata di semua dokumen, semakin rendah bobotnya.

## c) LDA

LDA adalah algoritma yang digunakan untuk menemukan topik-topik tersembunyi dalam kumpulan dokumen. Dalam langkah ini, model LDA dibuat dengan menentukan 10 topik menggunakan parameter `n_components=10`, yang berarti bahwa algoritma akan mencoba menemukan 10 topik berbeda dari data. Parameter `max_iter=10` menentukan bahwa model akan dilatih selama 10 iterasi, yang merupakan jumlah perulangan untuk menyempurnakan pengelompokan topik berdasarkan distribusi kata-kata dalam dokumen.

Dengan menggunakan matriks TF-IDF yang telah dibuat sebelumnya, model LDA di-fit untuk mengidentifikasi pola distribusi kata yang muncul bersama-sama dalam dokumen dan membentuk masing-masing dari 10 topik yang diprediksi. Hasil dari model ini nantinya akan berupa kumpulan topik yang diwakili oleh kumpulan kata-kata yang paling dominan atau sering muncul dalam dokumen terkait.

Setelah LDA dilakukan fungsi `display_topics` digunakan untuk mengekstrak dan menampilkan kata-kata utama dari setiap topik yang telah diidentifikasi oleh model LDA. Setelah model LDA dilatih pada data, ia menghasilkan sejumlah topik yang terdiri dari kata-kata yang sering muncul bersamaan. Fungsi ini mengiterasi melalui setiap topik yang dihasilkan oleh model dan menampilkan daftar kata-kata kunci dari masing-masing topik. Dalam kasus ini, untuk setiap topik, 10 kata kunci teratas (`no_top_words = 10`) yang memiliki bobot paling signifikan dalam membentuk topik tersebut akan ditampilkan. Kata-kata ini merupakan representasi dari isi atau tema topik yang ditemukan.

## d) Distribusi Topik dan Frekuensi Kemunculan Topik

Langkah berikutnya adalah identifikasi topik dominan di setiap dokumen serta penghitungan frekuensi kemunculan setiap topik. Distribusi topik dalam setiap dokumen dihitung menggunakan metode `transform` pada matriks TF-IDF, yang menghasilkan probabilitas keterkaitan dokumen dengan setiap topik. Dengan menggunakan fungsi `np.argmax`, topik dengan probabilitas tertinggi di setiap dokumen ditetapkan sebagai topik dominan. Kemudian frekuensi kemunculan setiap topik dihitung menggunakan `np.bincount` untuk mengetahui seberapa sering masing-masing topik muncul sebagai topik dominan di seluruh dokumen.

Urutan topik disusun berdasarkan frekuensi kemunculannya, di mana topik dengan jumlah kemunculan tertinggi dianggap sebagai topik yang paling sering muncul dan relevan. Hasil akhir dari proses ini adalah pengidentifikasian 10 topik yang paling sering muncul beserta frekuensi kemunculannya, yang kemudian ditampilkan dalam bentuk grafik batang untuk memberikan visualisasi dari topik trending di kumpulan dokumen.

## e) Trending Topic

Dari hasil LDA dihasilkan 10 topik yang diurutkan berdasarkan frekuensi kemunculan kata dalam setiap *text*. Adapun 10 topik tersebut adalah:

- 1) Topik 6 dengan kemunculan 86 kali  
Kata-kata kunci topik: **di ntb unggul akreditasi raih satusatunya pertama uin formasi yang**
- 2) Topik 4 dengan kemunculan 68 kali  
Kata-kata kunci topik: **mahasiswa ftk dengan upgrading metro buktikan islam tinggi perguruan universitas**
- 3) Topik 2 dengan kemunculan 50 kali  
Kata-kata kunci topik: **post dari banpt lombok selamat akreditasi unggul raih terima alumninya**
- 4) Topik 10 dengan kemunculan 48 kali  
Kata-kata kunci topik: **pilkada akademisi satu satunya kian nilai matang ntb predikat uin**
- 5) Topik 5 dengan kemunculan 48 kali  
Kata-kata kunci topik: **rinjani warta motogp parade padati puluhan pembalap ribu warga lombok**
- 6) Topik 8 dengan kemunculan 42 kali  
Kata-kata kunci topik: **akreditas raih unggul selamat uin dukung halal wisata tegaskan sinergi**
- 7) Topik 7 dengan kemunculan 38 kali  
Kata-kata kunci topik: **risiko utama kebebasan pilar demokrasi sipil sebagai mendorong spi manajemen**
- 8) Topik 1 dengan kemunculan 36 kali  
Kata-kata kunci topik: **disipliner diskusi dosen bersama lintas ftk gelar upsi malaysia metro**
- 9) Topik 9 dengan kemunculan 24 kali  
Kata-kata kunci topik: **ilmiah ptkin riset jawaban global tantangan terhadap selengkapnya oleh sejarah**
- 10) Topik 3 dengan kemunculan 20 kali  
Kata-kata kunci topik: **apresiasi indeks kepuasan peningkatan jemaah haji indonesia rektor metro ntb**

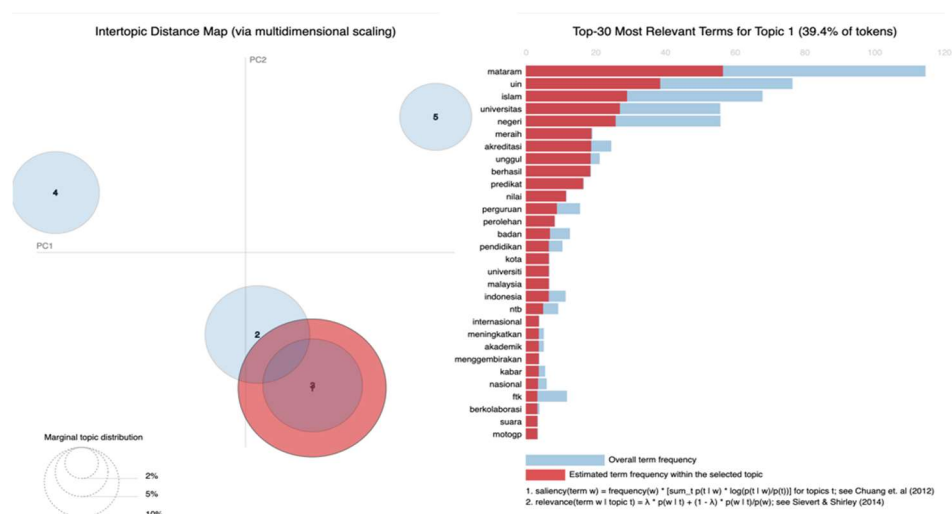
f) Intertopic Distance Map

*Intertopic Distance Map* (via *Multidimensional Scaling*) adalah salah satu fitur visualisasi yang disediakan oleh pyLDavis untuk membantu memahami hasil dari model *Latent Dirichlet Allocation* (LDA). Visualisasi ini menggambarkan hubungan antar-topik yang ditemukan oleh LDA, topik-topik tersebut berinteraksi dan berkaitan satu sama lain dalam ruang topik.

- 1) Peta Jarak Antar Topik (*Intertopic Distance Map*)
- 2) Pada peta ini, setiap topik yang ditemukan oleh LDA direpresentasikan sebagai lingkaran. Lingkaran-lingkaran ini ditempatkan di ruang dua dimensi (2D) menggunakan *Multidimensional Scaling* (MDS), yaitu metode untuk mereduksi dimensi sehingga kita bisa mengamati hubungan antar-topik dalam ruang yang lebih sederhana (2D).
- 3) Ukuran Lingkaran merupakan ukuran setiap lingkaran menunjukkan prevalensi relatif (dominasi atau proporsi) dari topik tersebut dalam dataset. Lingkaran yang lebih besar berarti topik tersebut muncul lebih sering di dokumen, sementara lingkaran yang lebih kecil berarti topik tersebut kurang dominan.
- 4) Jarak Antar Lingkaran adalah jarak antara dua lingkaran menunjukkan seberapa berbeda dua topik tersebut. Topik yang dekat satu sama lain (lingkarannya berdekatan) memiliki distribusi kata-kata yang serupa, artinya kata-kata yang muncul di kedua topik ini sering kali tumpang tindih. Topik yang jauh satu sama lain

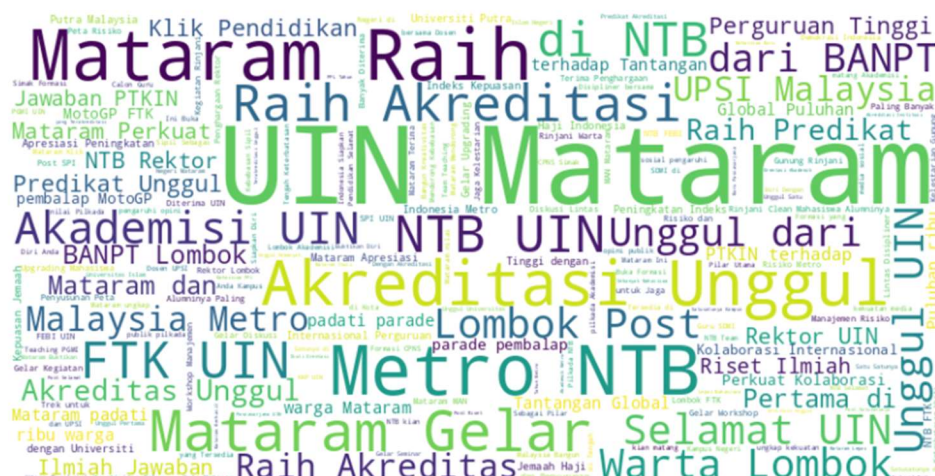
memiliki distribusi kata-kata yang berbeda, artinya topik-topik tersebut tidak banyak berbagi kata kunci yang sama.

- 5) Multidimensional Scaling (MDS) adalah teknik yang digunakan untuk memetakan hubungan antar-topik dari ruang dimensi yang lebih tinggi ke ruang dua dimensi (2D), yang kemudian divisualisasikan sebagai peta jarak antar topik. Tujuannya adalah untuk mereduksi dimensi topik (biasanya, distribusi kata di seluruh topik memiliki dimensi yang sangat tinggi) menjadi dua dimensi, sambil tetap mempertahankan hubungan relatif antara topik. Dalam peta ini, dimensi mewakili sifat-sifat yang sulit diinterpretasikan secara intuitif karena MDS tidak memetakan dimensi pada fitur tertentu, tetapi meminimalkan jarak antara topik dalam ruang dua dimensi.
- 6) Visualisasi, jika beberapa lingkaran saling berdekatan, ini berarti topik-topik tersebut cenderung memiliki kata-kata yang mirip dan sering berbagi konten yang sama. Ini bisa berarti topik-topik ini secara tematis berhubungan. Topik yang tersebar jauh menunjukkan bahwa mereka membahas topik yang berbeda dan memiliki kumpulan kata yang berbeda. Lingkaran yang lebih besar menunjukkan bahwa topik tersebut memiliki lebih banyak dokumen yang terkait dengannya, atau lebih sering muncul dalam dokumen-dokumen dataset.
- 7) Koherensi Topik, koherensi topik mengacu pada seberapa "masuk akal" sebuah topik berdasarkan kata-kata yang mengelompokkannya. Peta jarak ini membantu melihat apakah model LDA berhasil dalam mengelompokkan topik-topik yang memang saling berhubungan. Topik yang terlalu berdekatan bisa menjadi indikasi bahwa model mungkin terlalu banyak menempatkan topik-topik yang mirip, sehingga menyebabkan adanya redundansi topik. Sebaliknya, jika ada topik-topik yang sangat berjauhan tetapi berbagi banyak kata kunci, bisa jadi model LDA memisahkan topik-topik yang seharusnya terkait.
- 8) Hover Interaktif (*Interactive Hovering*), pada visualisasi pyLDavis yang berbentuk *hover* (melayang) di atas lingkaran untuk melihat kata-kata paling umum yang terkait dengan setiap topik. Ini membantu untuk mengidentifikasi apakah kata-kata yang ditemukan oleh LDA relevan dengan topik tersebut. Kata-kata kunci yang muncul di setiap topik memberikan gambaran tentang apa yang sebenarnya dibahas oleh topik tersebut. Hasil Intertopic Distance Map ditunjukkan pada Gambar 4, sementara hasil visualisasi dalam bentuk wordcloud ditunjukkan pada Gambar 5. Dimana pada wordcloud semakin besar ukuran kata dalam wordcloud menunjukkan semakin besar frekuensi kata tersebut muncul dalam dataset.



Gambar 4. Intertopic Distance Map





**Gambar 5. Wordcloud**

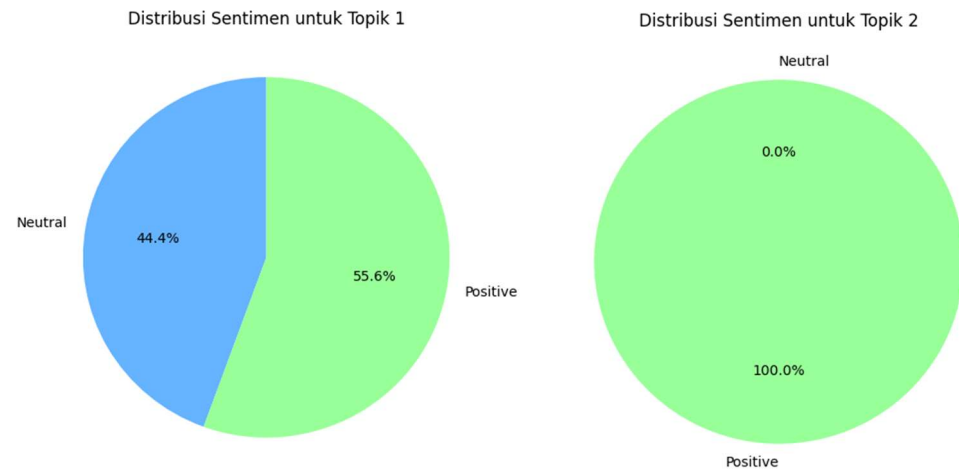
#### 4.2. Sentiment Analysis

Setelah trending topik ditemukan dengan menggunakan LDA, program akan memuat kamus kata sentimen. Langkah ini mempersiapkan sumber daya untuk analisis sentimen. Kamus kata positif dan negatif yang digunakan untuk analisis sentimen disimpan dalam file 'positive-words.txt' dan 'negative-words.txt' yang disimpan di dalam *Google Drive*. Kamus ini berisi daftar kata-kata yang masing-masing dikategorikan sebagai positif atau negatif. Kamus ini digunakan untuk mengevaluasi kata-kata dari setiap dokumen dan menentukan sentimennya. Metode penentuan sentimen dengan kamus kata ini disebut dengan *lexicon based method*.

Proses sentiment analysis pada trending topik yang dihasilkan oleh LDA melibatkan beberapa tahapan. Pertama, setelah topik-topik diidentifikasi menggunakan *Latent Dirichlet Allocation* (LDA), setiap topik diwakili oleh kumpulan kata-kata utama yang paling sering muncul.

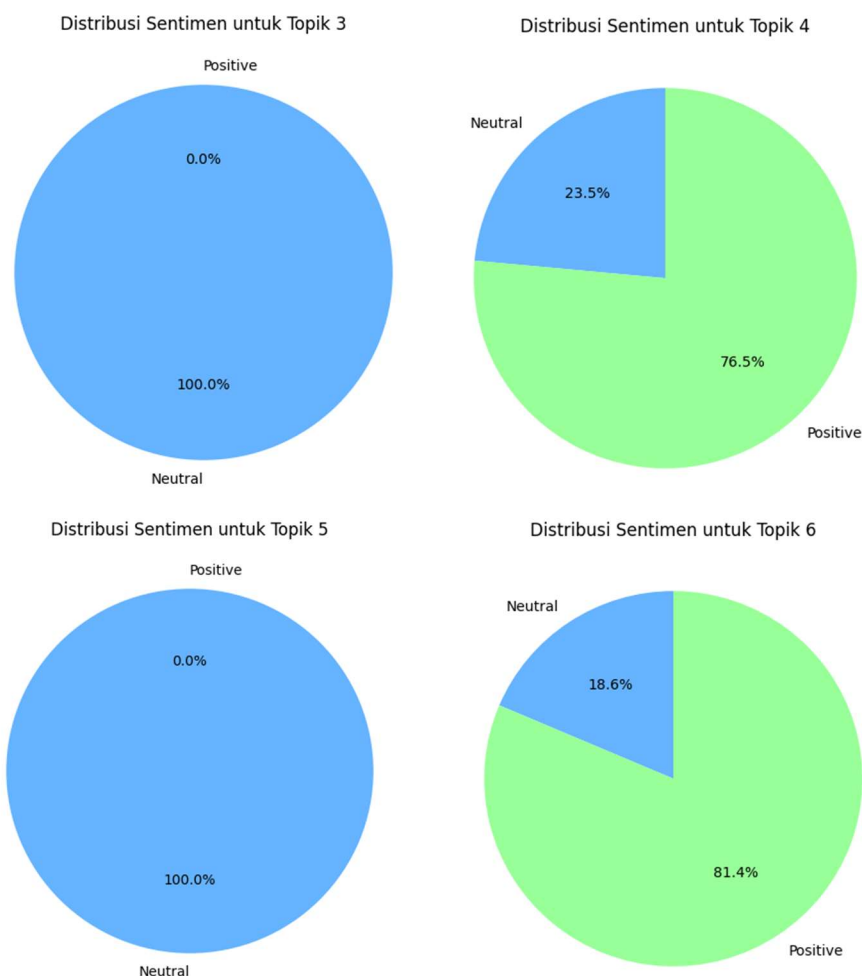
Untuk memahami sentimen dari setiap topik, kita melakukan analisis sentimen menggunakan metode lexicon-based dengan memanfaatkan kamus kata positif dan negatif. Kata-kata dari setiap dokumen yang terkait dengan topik dianalisis dan dibandingkan dengan daftar kata positif (disimpan dalam file `positive-words.txt`) dan kata negatif (disimpan dalam file `negative-words.txt`). Setiap kali sebuah kata ditemukan dalam salah satu dari kamus tersebut, itu akan berkontribusi pada skor sentimen yang positif atau negatif. Setelah sentimen dari semua dokumen yang berkaitan dengan sebuah topik dianalisis, kita dapat menghitung persentase dokumen yang memiliki sentimen positif atau negatif.

Hasil akhir dari analisis ini ditampilkan dalam bentuk diagram lingkaran, yang menunjukkan distribusi sentimen untuk setiap topik trending. Ini membantu memahami apakah topik tersebut cenderung menerima respons positif, negatif, atau netral dari audiens di media online. Analisis ini penting karena memberikan wawasan mendalam tentang bagaimana perasaan publik terhadap topik-topik yang sedang populer. Hasil sentiment untuk trending topik 1 sampai dengan trending topik 2 ditunjukkan pada Gambar 6.



**Gambar 6.** Hasil Sentiment Trending 1 dan 2

Hasil sentiment untuk trending topik 3, trending 4, trending 5 dan 6 ditunjukkan pada Gambar 7. Berdasarkan hasil analisis sentimen yang diterapkan pada 10 topik utama yang ditemukan melalui LDA, terlihat bahwa sebagian besar topik memiliki sentimen yang cenderung positif, dengan beberapa variasi dalam proporsi sentimen netral. Topik yang paling dominan, yaitu Topik 6, memiliki proporsi 81.4% sentimen positif, yang menunjukkan bahwa diskusi mengenai akreditasi unggul dan pencapaian UIN di NTB umumnya disambut baik oleh audiens.

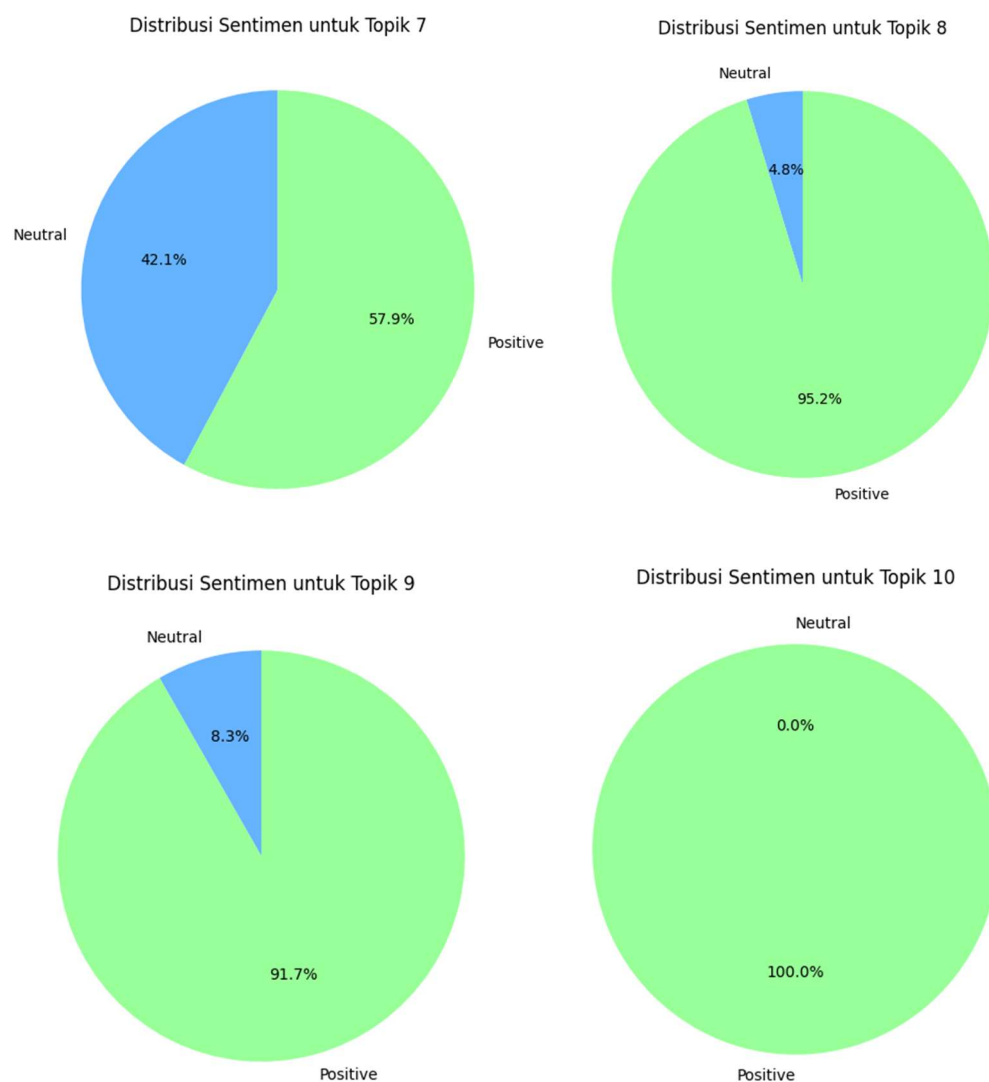


**Gambar 7.** *Sentiment Trending Topik 3, 4, 5 dan 6*

Topik 8 juga menonjol dengan 95.2% sentimen positif, menunjukkan bahwa dukungan terhadap wisata halal dan sinergi UIN mendapatkan respon yang sangat baik. Hasil sentiment trending topik 7, 8, 9 & 10 ditunjukkan pada Gambar 8.

Di sisi lain, ada beberapa topik yang menunjukkan sentimen netral yang lebih tinggi. Misalnya, Topik 3 dan Topik 5 memiliki 100% sentimen netral, yang menunjukkan bahwa diskusi di sekitar riset ilmiah dan parade MotoGP di Lombok lebih bersifat informatif tanpa menunjukkan emosi yang kuat.

Hal ini mungkin dikarenakan sifat topik yang lebih bersifat deskriptif atau kurang memicu emosi tertentu. Topik 2, Topik 10, dan Topik 9 menampilkan 100% dan 91.7% sentimen positif, mengindikasikan bahwa berita tentang pencapaian akreditasi, nilai UIN, serta tantangan riset dan sejarah mendapatkan respon yang sangat positif dari pembaca.



**Gambar 8.** *Sentiment Trending 7, 8, 9, dan 10*

Keseluruhan analisis sentimen menunjukkan bahwa UIN Mataram dan kegiatan-kegiatan yang terkait dengannya, seperti akreditasi unggul, riset ilmiah, dan wisata halal, memberikan dampak positif yang signifikan di kalangan publik. Namun, ada beberapa topik yang mungkin dianggap lebih netral, seperti diskusi tentang riset ilmiah atau parade MotoGP, yang menunjukkan bahwa topik-topik ini tidak terlalu memancing reaksi emosional yang kuat.

## 5. Kesimpulan

Berdasarkan hasil implementasi penelitian yang sudah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Hasil pencarian trending dengan LDA, ditemukan 10 topik utama dari berita-berita terkait UIN Mataram dalam 60 hari terakhir. Topik-topik yang paling sering muncul antara lain terkait akreditasi unggul, prestasi mahasiswa, pencapaian akademik, serta event penting di NTB seperti parade MotoGP. Topik dengan kemunculan terbanyak adalah "Topik 6" yang terkait dengan akreditasi unggul di NTB, diikuti oleh "Topik 4" tentang pencapaian mahasiswa di universitas. Topik-topik ini menunjukkan fokus pada pencapaian dan kontribusi UIN Mataram serta isu-isu lokal di NTB.
2. Dari analisis sentimen menggunakan metode lexicon-based, sebagian besar topik terkait UIN Mataram mendapatkan sentimen positif yang dominan. Beberapa topik

menunjukkan sentimen positif hampir mencapai 100%, seperti "Topik 2" (100% positif) yang berkaitan dengan pencapaian akreditasi, dan "Topik 10" yang juga sepenuhnya positif. Sementara itu, beberapa topik lain, seperti "Topik 3" dan "Topik 5", lebih didominasi oleh sentimen netral, yang menunjukkan bahwa diskusi tentang topik tersebut cenderung bersifat informatif tanpa memicu emosi yang signifikan. Secara keseluruhan, institusi UIN Mataram dilihat secara positif dalam berita dan diskusi media yang dianalisis.

## Referensi

- [1] I. N. Agustiani, R. S. Y. Zebua, M. R. P. Kusyanda, N. I. Rusdiani, N. Hayani, and A. A. Rizal, *BUKU AJAR DIGITAL MARKETING*, vol. 1. Jambi: Sonpedia, 2024. <https://buku.sonpedia.com/2024/02/buku-ajar-digital-marketing.html>
- [2] B. O. Tafakkur, L. P. I. Kharisma, A. A. Rizal, and A. Abdurahim, "Implementasi Augmented Reality Sebagai Media Promosi Pada Lesehan Kalisari Dengan Metode Based Marker Tracker," *JTIM: Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 1, pp. 10–21, May 2023, doi: <https://doi.org/10.35746/jtim.v5i1.331>.
- [3] V. Chakkarwar and S. Tamane, "Social Media Analytics during Pandemic for Covid19 using Topic Modeling," in *Proceedings of the 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing, ICSIDEMPC 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 279–282. doi: <https://doi.org/10.1109/ICSIDEMPC49020.2020.9299617>.
- [4] Q. Wu and L. He, "Analysis of Social Media User Network Access Behavior Based on Web4.0," in *2019 6th International Conference on Dependable Systems and Their Applications (DSA)*, Institute of Electrical and Electronics Engineers Inc., Mar. 2020, pp. 127–134. doi: <https://doi.org/10.1109/DSA.2019.00023>.
- [5] P. Jain and K. Y. Chan, "A Social Media Analytical Framework Incorporating Fuzzy Regression for Affective Design," in *2020 International Conference on System Science and Engineering (ICSSE)*, IEEE, Oct. 2020. doi: <https://doi.org/10.1109/ICSSE50014.2020.9219261>.
- [6] S. Junaidi et al., *BUKU AJAR MACHINE LEARNING*. Jambi: Sonpedia, 2024. <https://buku.sonpedia.com/2024/02/buku-ajar-machine-learning.html>
- [7] N. Amy, A. Muchali, J. Budiarto, A. B. Maulachela, and A. A. Rizal, "Management of issue and public trust using information technology," *J Phys Conf Ser*, vol. 1539, no. 1, 2020, doi: <https://doi.org/10.1088/1742-6596/1539/1/012073>.
- [8] Y. Chen, H. Zhang, R. Liu, Z. Ye, and J. Lin, "Experimental explorations on short text topic mining between LDA and NMF based Schemes," *Knowl Based Syst*, vol. 163, pp. 1–13, Jan. 2019, doi: <https://doi.org/10.1016/j.knosys.2018.08.011>.
- [9] A. R. Alharbi, S. D. Alharbi, A. Aljaedi, and O. Akanbi, "Neural Networks Based on Latent Dirichlet Allocation for News Web Page Classifications," in *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020. doi: <https://doi.org/10.1109/IICAJET49801.2020.9257842>.
- [10] A. A. Rizal, G. S. Nugraha, R. A. Putra, and D. P. Anggraeni, "Twitter Sentiment Analysis in Tourism with Polynomial Naïve Bayes Classifier," *JTIM: Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 4, pp. 343–353, Feb. 2024, doi: <https://doi.org/10.35746/jtim.v5i4.478>.
- [11] A. Priadana and A. A. Rizal, "Sentiment Analysis on Government Performance in Tourism During The COVID-19 Pandemic Period With Lexicon Based," *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, vol. 7, no. 1, pp. 28–39, Nov. 2021, doi: <https://doi.org/10.18860/ca.v7i1.12488>.
- [12] J. Cha, S. Kim, and E. Park, "A lexicon-based approach to examine depression detection in social media: the case of Twitter and university community," *Humanit Soc Sci Commun*, vol. 9, no. 1, Dec. 2022, doi: <https://doi.org/10.1057/s41599-022-01313-2>.
- [13] M. N. Asti, A. A. Rizal, and I. Ismarmiaty, "Lexicon Based Sentiment Analysis pada Trending Topics di Nusa Tenggara Barat," *JICOM: Jurnal Informatika dan Komputer*, vol. 3, no. 2, pp. 93–98, 2022, <https://ejurnalunsam.id/index.php/jicom/article/view/6136>.
- [14] L. Antón-González, M. Pans, J. Devís-Devís, and L. M. González, "Cycling in urban environments: Quantitative text analysis," *J Transp Health*, vol. 32, Sep. 2023, doi: <https://doi.org/10.1016/j.jth.2023.101651>.
- [15] S. Rifky et al., *ARTIFICIAL INTELLIGENCE (Teori dan Penerapan AI di Berbagai Bidang)*. Jambi: PT. Sonpedia Publishing Indonesia, 2024, <https://buku.sonpedia.com/2024/05/artificial-intelligence.html>.