

## ANALISIS ISU SOSIAL MAHASISWA BERBASIS DATA MEDIA SOSIAL MENGUNAKAN *LATENT DIRICHLET ALLOCATION* (LDA)

### [Analysis Of Student Social Issues Based On Social Media Data Using Latent Dirichlet Allocation (LDA)]

Jesisca<sup>1)</sup>, Chodri Dwi Yonny<sup>2)\*</sup>

Teknologi Informasi, Universitas Bangka Belitung, Balun Ijuk, Indonesia

<sup>1)</sup>*jesisca052@gmail.com*, <sup>2)</sup>*chodriyonny37@gmail.com (corresponding)*

#### ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi isu-isu sosial yang paling sering dibahas oleh mahasiswa di media sosial seperti TikTok dan Instagram. Data dikumpulkan melalui kuesioner yang disebarakan kepada 50 mahasiswa, dengan 49 data valid yang dianalisis. Metode Latent Dirichlet Allocation (LDA) berbasis aljabar linier digunakan untuk mengelompokkan teks ke dalam beberapa topik utama.

Hasil analisis mengungkapkan tiga klaster isu sosial utama, yaitu (1) kesetaraan gender dan diskriminasi lingkungan kampus, (2) biaya pendidikan, kesehatan mental, dan kesejahteraan mahasiswa, serta (3) politik kampus dan peluang karier. Topik mengenai politik kampus merupakan yang paling dominan di antara yang lain. Temuan ini menunjukkan bahwa mahasiswa aktif menggunakan media sosial untuk menyampaikan pendapat dan menanggapi kebijakan kampus yang berdampak langsung pada kehidupan akademik mereka. Selain itu, penelitian ini membuktikan efektivitas metode LDA dalam menganalisis data sosial untuk memahami pola komunikasi digital di kalangan mahasiswa.

---

**Kata kunci:** LDA; isu sosial mahasiswa; media sosial; aljabar linier; analisis teks

#### ABSTRACT

*This study aims to identify the most frequently discussed social issues among university student on social media platforms such as TikTok and Instagram. Data were collected via a questionnaire distributed to 50 students, with 49 valid responses analyzed. The Latent Dirichlet Allocation (LDA) method based on linear algebra was applied to classify text data into several main topics.*

*The result revealed three primary clusters of social issues: (1) gender equality and discrimination within the campus environment, (2) educational costs, mental health, and student welfare, and (3) campus politics and career opportunities. Campus politics emerged as the most dominant topic among others. These findings suggest that student actively utilize social media to express their views and respond to institutional policies that directly affect their academic lives. Furthermore, this research demonstrates the effectiveness of the LDA method in analyzing social data to reveal digital communication patterns among student.*

---

**Keywords:** LDA; student social issues; social media; linear algebra; text analysis

#### PENDAHULUAN

Perkembangan media sosial di era digital telah mengubah cara mahasiswa berinteraksi dan menyampaikan isu-isu sosial. Platform seperti TikTok dan Instagram bukan hanya untuk media hiburan saja, tetapi juga berfungsi sebagai ruang mahasiswa untuk mengekspresikan pendapat, berdiskusi melalui konten-konten viral, dan membangun wacana sosial lewat penggunaan hashtag. Hal ini menunjukkan bahwa mahasiswa semakin aktif di ranah publik digital, berperan lebih dari sekedar konsumen informasi menjadi pencipta wacana sosial. Menurut (Azzura et al., 2025), media sosial berperan penting dalam mengungkap berbagai permasalahan akademik dan sosial yang

dihadapi mahasiswa, serta menjadi wadah untuk menyuarakan pandangan terkait isu-isu yang berkembang di lingkungan kampus dan masyarakat.

Namun, dalam jumlah besar dan beragamnya konten yang dibuat oleh mahasiswa di media sosial membuat metode analisis tradisional menjadi kurang efektif. Oleh karena itu, pendekatan menggunakan machine learning seperti *Latent Dirichlet Allocation (LDA)* diperlukan untuk menggali pola tema tersembunyi dari kumpulan teks digital yang besar. Menurut (Surya Buana et al., 2024) metode pemodelan topik berbasis LDA mampu menganalisis konten media sosial dan mengidentifikasi tema-tema utama yang dibahas oleh pengguna. Model ini mampu secara otomatis mengidentifikasi topik utama dari ratusan hingga ribuan unggahan, memberikan gambaran objektif mengenai isu sosial yang paling banyak dibahas oleh mahasiswa di media sosial. Berdasarkan hal tersebut, penelitian ini fokus menganalisis isu sosial mahasiswa melalui data unggahan berbasis hashtag di Tiktok dan Instagram dengan menggunakan metode LDA sebagai alat analisis utama. (Puspita et al., 2024) juga menyatakan bahwa LDA bisa diterapkan tidak hanya pada media sosial, tetapi juga pada berbagai sumber data daring lainnya seperti portal berita.

Tujuan penelitian ini adalah untuk menemukan tema utama yang muncul dari aktivitas sosial dan ekspresi mahasiswa di media sosial, menganalisis tren kemunculan isu-isu tersebut, serta menafsirkan konteks sosial yang terkait. Lingkup penelitian mencakup data unggahan dengan hashtag di TikTok dan Instagram. Tahapan analisis meliputi pemrosesan awal teks, penerapan model LDA, serta interpretasi hasil berdasarkan kata kunci dominan dalam topik. Secara akademik, penelitian ini diharapkan dapat memberikan kontribusi pada kajian analisis wacana digital berbasis data khususnya dalam konteks mahasiswa. Secara praktis, hasil penelitian dapat membantu institusi Pendidikan dan organisasi kemahasiswaan untuk lebih memahami isu sosial yang sedang hangat di ranah digital.

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan *text mining* untuk menganalisis isu sosial yang sedang hangat di lingkungan mahasiswa di media sosial. Proses analisis melibatkan ekstraksi data teks serta pemodelan topik menggunakan metode *Latent Dirichlet Allocation (LDA)*, salah satu teknik topic modelling terkemuka dalam pengolahan teks digital (Nawang Sari & Dwi Purnomo, 2022). Data yang digunakan berasal dari dua sumber, yaitu hasil kuesioner yang diisi oleh 50 mahasiswa dengan 49 data valid dan data observasi media sosial berupa hashtag relevan. Seluruh data dikompilasi dalam format Excel dan diolah melalui lingkungan pemrograman Google Colab.

Analisis dilakukan melalui tahapan *preprocessing* data teks meliputi penghapusan tanda baca, konversi teks ke huruf kecil, dan eliminasi stopwords. Selanjutnya, teks direpresentasikan ke dalam matrix dokumen-kata menggunakan teknik *CountVectorizer*. Metode LDA diterapkan untuk mengidentifikasi distribusi topik laten dalam dokumen, dengan hasil akhir divisualisasikan menggunakan word cloud (Andika et al., 2021). Pendekatan ini didasarkan pada kerangka kerja yang serupa dalam studi topic modelling pada media sosial di Indonesia (Alpiana et al., 2024).

Metode LDA menguraikan matriks dokumen-kata menjadi dua matriks baru : matriks dokumen-topik dan matriks dokumen-kata . Proses ini merupakan implementasi dari konsep aljabar linier, dimana representasi teks difaktorkan secara linear (Jelodar et al., 2017). Penentuan jumlah topik sebanyak tiga topik dilakukan untuk merepresentasikan isu sosial utama yang berkembang di kalangan mahasiswa. Proses komputasi dan analisis menggunakan bahasa pemrograman *Python* dengan library pendukung antara lain *pandas*, *scikit-learn*, *matplotlib*, serta *Sastrawi*. Interpretasi hasil dilakukan berdasarkan bobot frekuensi kata dalam setiap topik dan distribusi topik pada masing-masing responden kuesioner.

## HASIL DAN PEMBAHASAN

Hasil analisis menggunakan metode *Latent Dirichlet Allocation (LDA)* pada 49 data valid responden mahasiswa mengenai isu sosial yang sering mereka temui di media sosial (TikTok dan

Instagram). Model LDA menghasilkan tiga topik utama yang mewakili kelompok isu sosial mahasiswa.

### Hasil Visualisasi WordCloud

Visualisasi *WordCloud* digunakan untuk menampilkan kata yang paling sering muncul pada tiap topik. Semakin besar ukuran kata, semakin tinggi bobot kontribusinya terhadap topik (Nafi et al., 2024).



**Gambar 1. Word Cloud Topik 1**  
(Sumber : Hasil Analisis Python,2025)



**Gambar 2. Word Cloud Topik 2**  
(Sumber : Hasil Analisis Python,2025)



**Gambar 3. Word Cloud Topik**  
(Sumber : Hasil Analisis Python,2025)

Interpretasi WordCloud :

- Topik 1 didominasi oleh kata *diskriminasi*, *gender*, *kesetaraan*.
- Topik 2 didominasi oleh kata *pendidikan*, *biaya*, *UKT*, *mental*.
- Topik 3 didominasi oleh kata *politik*, *negara*, *kampus*.

### Matriks Dokumen-Topik

LDA menghasilkan matriks berukuran 49 x 3, di mana setiap nilai menunjukkan probabilitas suatu dokumen termasuk ke dalam sebuah topik.

**Tabel 1. Distribusi Probabilitas Dokumen ke Topik**

| No | Teks Responden                         | T1           | T2           | T3           | Topik dominan |
|----|--|--------------|--------------|--------------|---------------|
| 1  | Politik kampus/<br>negara              | 0,083        | 0,083        | <b>0,833</b> | Topik 3       |
| 2  | Kesehatan mental<br>mahasiswa          | 0,084        | <b>0,833</b> | 0,083        | Topik 2       |
| 3  | Kesetaraan<br>gender &<br>diskriminasi | <b>0,833</b> | 0,083        | 0,083        | Topik 1       |
| 4  | Perubahan iklim<br>& lingkungan        | 0,084        | 0,084        | <b>0,832</b> | Topik 3       |
| 5  | Biaya Pendidikan<br>& UKT              | 0,083        | <b>0,833</b> | 0,083        | Topik 2       |

(Sumber : Hasil Analisis Python,2025)

Matriks dokumen-topik berukuran 49 x 3, yang merepresentasikan 49 data responden terhadap tiga topik utama hasil pemodelan LDA.

Hasil distribusi topik :

- Topik 1 = 10 responden (20%)
- Topik 2 = 15 responden (30%)
- Topik 3 = 24 responden (50%)

Hal ini menunjukkan bahwa isu politik kampus paling sering muncul pada mahasiswa.

### Matriks Topik-Kata

**Tabel 2. Kata Dominan dan Fokus isu Hasil LDA**

| Topik | 5 Kata Dominan                                       | Fokus Isu                           |
|-------|--|-------------------------------------|
| T1    | Diskriminasi, gender, kesetaraan, kampus, lingkungan | Kesetaraan & hak mahasiswa          |
| T2    | Biaya, UKT, pendidikan, kesehatan, mental            | Kesejahteraan & finansial mahasiswa |
| T3    | Politik, negara, kampus, magang, kerja               | Politik kampus & peluang kerja      |

(Sumber : Hasil Analisis Python,2025)

Dari tabel di atas, dapat dilihat bahwa sumber utama permasalahan sosial yang sering dibicarakan oleh mahasiswa dominan terdapat kata “kampus” pada setiap topik.

### Pembahasan Hasil

Analisis data menunjukkan bahwa mahasiswa secara aktif memanfaatkan media sosial sebagai ruang diskusi untuk mengekspresikan keresahan dan pandangan mereka terhadap berbagai isu sosial yang mempengaruhi kehidupan akademik. Berdasarkan distribusi probabilitas dari metode LDA, topik ketiga yang membahas politik kampus dan kebijakan pendidikan menjadi yang paling dominan dengan persentase mencapai 50%. Temuan ini mengindikasikan Tingkat kepedulian yang tinggi dari mahasiswa terhadap arah dan dampak kebijakan Pendidikan nasional, termasuk terkait peluang kerja, program magang, dan otonomi institusi kampus.

Keterkaitan erat isu kampus dengan ranah politik dan kebijakan politik menegaskan bahwa mahasiswa saat ini berperan sebagai agen aktif, bukan sekadar pengamat dalam proses perubahan kebijakan pendidikan. Media sosial berfungsi sebagai media advokasi sekaligus alat pengawasan terhadap perubahan kebijakan yang berpotensi merugikan kepentingan mahasiswa. Selanjutnya, topik kedua, dengan porsi sebesar 30%, memfokuskan pada isu biaya pendidikan, UKT, serta kesejahteraan mental mahasiswa. Hal ini memperlihatkan bahwa beban ekonomi dan tekanan psikologis masih menjadi tantangan utama dalam sistem pendidikan tinggi. Berbagai diskusi mengemuka mengenai kenaikan biaya kuliah, tekanan akademik, burnout, serta minimnya akses terhadap layanan konseling yang memadai (Wilkes & Riswani, 2022).

Topik pertama, meski memiliki proporsi paling kecil sebesar 20%, menyoroti isu kesetaraan gender, diskriminasi, dan pengembangan lingkungan kampus yang inklusif. Keberadaan isu tetap

penting karena mencerminkan permasalahan ketidakadilan struktural yang masih perlu mendapat perhatian dan perjuangan terus menerus (Farqi, 2023).

Secara keseluruhan, isu yang bersifat emosional dan berdampak langsung pada kesejahteraan mahasiswa lebih banyak dibahas dibanding topik yang jarang terkait dengan kehidupan kampus. Temuan ini memperkuat pandangan bahwa media sosial bukan sekadar platform hiburan, melainkan juga arena partisipasi aktif mahasiswa dalam berbagai problem sosial dan akademik.

### **Keterkaitan Temuan dengan Penelitian Sebelumnya**

Temuan penelitian ini memperkuat hasil-hasil studi terdahulu yang membahas partisipasi mahasiswa dalam dikursus isu sosial melalui media sosial. Nawang Sari & Dwi Purnomo, (2022) menyatakan bahwa mahasiswa aktif menggunakan platform digital seperti TikTok dan Instagram untuk menyampaikan isu-isu kampus, khususnya yang terkait dengan hasil penelitian, di mana isu politik kampus dan negara mendominasi pembahasan mahasiswa.

Selain itu, Buana et al. (2024) menekankan bahwa kenaikan biaya pendidikan, tekanan akademik, dan isu kesehatan mental merupakan tema sentral yang banyak dibahas mahasiswa di media sosial. Hal ini didukung oleh hasil LDA dalam penelitian ini yang menunjukkan bahwa topik mengenai UKT, kesehatan mental, dan beban akademik memiliki proporsi terbesar kedua dalam diskusi responden.

Selanjutnya, Puspita an Wulandari (2024) menemukan bahwa isu diskriminasi dan kesetaraan gender tetap menjadi perhatian mahasiswa dalam konteks pendidikan tinggi. Penelitian ini juga mengkonfirmasi keberadaan isu tersebut melalui identifikasi topik yang terkait dengan kesetaraan, diskriminasi, dan lingkungan kampus yang inklusif.

Lebih jauh, Azzura et al. (2025) membuktikan bahwa metode *Latent Dirichlet Allocation (LDA)* efektif dalam mengelompokkan opini publik dari platform digital secara otomatis berdasarkan pola tematik. Relevansi ini terlihat dalam penelitian saat ini, di mana LDA berhasil mengidentifikasi tiga isu utama dengan representasi probabilitas yang jelas dari data responden mahasiswa. Dengan demikian, hasil penelitian ini tidak hanya selaras dengan studi sebelumnya, tetapi juga memberikan kontribusi baru dalam konteks analisis isu sosial mahasiswa melalui pendekatan *text mining* berbasis LDA.

## **PENUTUP**

### **Simpulan**

Penelitian ini bertujuan untuk menganalisis isu sosial yang paling sering dibahas oleh mahasiswa melalui media sosial, khususnya Instagram dan TikTok, dengan menggunakan metode *Latent Dirichlet Allocation (LDA)*. Dari pemodelan topik yang dilakukan pada 49 data valid, ditemukan tiga kelompok isu utama, isu politik kampus menjadi yang paling dominan. Hasil ini menunjukkan bahwa media sosial berperan penting sebagai wadah bagi mahasiswa untuk mengungkapkan pandangan mereka terhadap isu-isu yang memengaruhi kehidupan akademik mereka. Selain itu, LDA terbukti sebagai metode yang efektif dalam menganalisis teks dan mengidentifikasi pola pembahasan mahasiswa di dunia digital.

### **Saran**

Penelitian berikutnya disarankan untuk memperluas jumlah data dengan melibatkan lebih banyak platform media sosial agar cakupan isu yang ditemukan menjadi lebih luas dan komprehensif. Selain itu, pengembangan metode analisis yang lebih canggih, seperti model pemodelan topik berbasis *transformer* atau pendekatan *deep learning*, dapat memberikan hasil yang lebih tepat. Institusi pendidikan dan organisasi kemahasiswaan juga diharapkan dapat menggunakan temuan ini sebagai dasar dalam merancang kebijakan dan program yang lebih responsif terhadap kebutuhan dan kekhawatiran mahasiswa.



## DAFTAR PUSTAKA

- Alpiana, V., Salam, A., Alzami, F., Rizqa, I., & Aqmala, D. (2024). Analisis Topic-Modelling Menggunakan Latent Dirichlet Allocation (LDA) Pada Ulasan Sosial Media Youtube. *Jurnal Media Informatika Budidarma*, 8(1), 332. <https://doi.org/10.30865/mib.v8i1.7127>
- Andika, M., Chaerani, L., Data, K. K., Allocation, L. D., Online, M., & Topik, P. (2021). *Pemodelan Topik Berita pada Portal Berita Online Berbahasa Indonesia Menggunakan Latent Dirichlet Allocation (LDA)*. 20, 173–180.
- Azzura, E. P., Mutahir, A., Hariyadi, H., & Dadan, S. (2025). Penggunaan Bahasa dalam Media Sosial X untuk Mengungkapkan Permasalahan Akademik Mahasiswa (Analisis Wacana Kritis Akun @unsoedmfs). *Jurnal Penelitian Inovatif*, 5(3), 2205–2224. <https://doi.org/10.54082/jupin.566>
- Farqi, A. F. (2023). *Kesetaraan Gender Di Area Kampus Universitas Jember*. 3(4).
- Jelodar, H., Wang, Y., Yuan, C., & Feng, X. (2017). *Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey*.
- Nafi, K., Sholehah, R., & Rahim, N. (2024). *Unveiling Insights : A Knowledge Discovery Approach to Comparing Topic Modeling Techniques in Digital Health Research*. 8(1), 108–121.
- Nawang Sari, W. A., & Dwi Purnomo, H. (2022). Topic Modeling Using the Latent Dirichlet Allocation Method on Wikipedia Pandemic Covid-19 Data in Indonesia. *Jurnal Teknik Informatika (Jutif)*, 3(5), 1223–1230. <https://doi.org/10.20884/1.jutif.2022.3.5.321>
- Puspita, E., Shiddieq, D. F., & Roji, F. F. (2024). Topic Modeling on Online News Media Using Latent Diriclet Allocation (Case Study Somethinc Brand). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 481–489.
- Surya Buana, G., Tyasnurita, R., Cahya Puspita, N., Aulia Vinarti, R., & Mahananto, F. (2024). 1636-6477-1-Pb. *International Journal on Informatics Visualization*, 8(March), 88–95.
- Wilkes, J., & Riswani, R. (2022). *Mental Wellbeing of Indonesian Students : Mean Comparison with UK Students and Relationships with Self-Compassion and Academic Engagement*.