



TOPIC MODELLING POLA KOMUNIKASI PILPRES 2024 FOCUS WEB SCRAPING DAN LATENT DIRICHLET ALLOCATION

Tedi Irawan¹, Lalu Mutawalli^{2(*)}, Sofiansyah Fadli³, Wire Bagye³

^{1,3}Program Studi Teknik Informatika, STMIK Lombok,

^{2(*)}Program Studi Sistem Informasi, STMIK Lombok

Jln. Basuki Rahmat No.105 Praya Lombok Tengah 83511

¹tediirawan692@gmail.com, ²lualualistilo@gmail.com, ³wirestmik@gmail.com

Abstract

The 2024 Presidential Election in Indonesia is highly anticipated as it determines the direction of the nation's leadership. Candidates Anies-Muhaimin, Prabowo-Gibran, and Ganjar-Mahfud are competing closely. Digital communication strategies are crucial in the campaigns, leveraging social media to influence public opinion. Different approaches are adopted: Prabowo's "Gemoy" campaign, Anies-Muhaimin's door-to-door strategy, and Ganjar-Mahfud's focus on environmental issues. This research focuses on analyzing communication patterns in the context of the 2024 Presidential Election. Using web scraping and Latent Dirichlet Allocation (LDA) for topic modeling, this study aims to understand the dynamics of political communication in the digital era. It seeks to evaluate the coherence of the LDA model in identifying dominant topics discussed about the 2024 Presidential Election. Web scraping collects data from three datasets: AMIN, Prabowo-Gibran, and Ganjar-Mahfud. LDA is then applied for topic modeling of the extracted data. This research is novel in combining web scraping and LDA for political communication analysis in the digital age. The analysis shows that the LDA model has a coherence score of 0.405, which, although not very high, still provides valuable insights into the dominant topics discussed. These topics include candidate support, survey results, and political issues related to the 2024 Presidential Election. This research offers important insights into the dynamics of political communication in the digital realm, particularly in the context of the 2024 Presidential Election. Various parties, including candidates, political parties, and the general public, can use the information obtained to develop more effective communication strategies and understand the most relevant issues that may influence the election outcome.

Keywords : Topics; Model; Pilpres 2024; Web scraping; LDA.

Abstrak

Pemilihan Presiden 2024 di Indonesia sangat dinantikan karena menentukan arah kepemimpinan negara. Kandidat Anies-Muhaimin, Prabowo-Gibran, dan Ganjar-Mahfud bersaing ketat. Strategi komunikasi digital menjadi krusial dalam kampanye, memanfaatkan media sosial untuk mempengaruhi opini publik. Pendekatan berbeda diadopsi: kampanye "Gemoy" Prabowo, door-to-door Anies-Muhaimin, dan fokus isu lingkungan oleh Ganjar-Mahfud. Penelitian ini berfokus pada analisis pola komunikasi dalam konteks Pemilihan Presiden 2024. Dengan menggunakan metode web scraping dan Latent Dirichlet Allocation (LDA) untuk pemodelan topik, penelitian ini berusaha memahami dinamika komunikasi politik yang berkembang di era digital. Bagaimana koherensi model LDA dalam mengidentifikasi topik-topik, Mengidentifikasi topik-topik dominan yang muncul dalam diskusi terkait Pemilihan Presiden 2024. Penelitian ini menggunakan metode web scraping untuk mengumpulkan data dari tiga dataset berbeda: AMIN, Prabowo-Gibran, dan Ganjar-Mahfud. Metode LDA kemudian diterapkan untuk pemodelan topik dari data yang telah diekstraksi. Penelitian ini memiliki keterbaruan dalam pendekatannya yang menggunakan kombinasi web scraping dan LDA untuk analisis komunikasi politik di era digital. Hasil analisis menunjukkan bahwa model LDA memiliki tingkat koherensi sebesar 0,405, yang meskipun tidak sangat tinggi, tetap memberikan pemahaman berharga terhadap topik-topik dominan yang dibahas. Topik-topik tersebut meliputi dukungan terhadap kandidat, hasil survei, dan isu-isu politik terkait Pemilihan Presiden



2024. Penelitian ini memberikan wawasan penting dalam memahami dinamika komunikasi politik yang terjadi di ranah digital, khususnya dalam konteks Pemilihan Presiden 2024. Informasi yang diperoleh dapat digunakan oleh berbagai pihak, termasuk calon kandidat, partai politik, dan masyarakat umum, untuk mengembangkan strategi komunikasi yang lebih efektif dan memahami isu-isu yang paling relevan serta berpotensi memengaruhi hasil pemilihan.

Kata kunci : Topik, Model, Pilpres 2024, *Web scraping*, LDA.

1. PENDAHULUAN

Pemilihan Umum Presiden (Pilpres) adalah salah satu momen penting dalam kehidupan demokrasi di suatu negara. Di Indonesia, Pilpres menjadi peristiwa yang sangat dinantikan oleh masyarakat karena menentukan arah kepemimpinan negara untuk periode berikutnya. Komunikasi politik memiliki peran penting dalam menentukan hasil Pilpres, karena melalui komunikasi ini, para kandidat berusaha mempengaruhi opini dan preferensi pemilih[1][2].

Kemajuan teknologi informasi dan komunikasi, terutama internet dan media sosial, telah mengubah dinamika politik di Indonesia. Pada Pilpres sebelumnya, platform media sosial seperti Twitter, Facebook, dan Instagram telah menjadi sarana utama bagi kandidat dan tim kampanye untuk berinteraksi dengan pemilih potensial. Dalam konteks Pilpres 2024, dengan penetrasi internet dan penggunaan media sosial yang terus meningkat di Indonesia, peran media daring dalam membentuk opini publik dan memengaruhi preferensi pemilih menjadi lebih penting[3][4][5].

Dalam Pemilu Pilpres 2024, terdapat tiga pasangan calon presiden dan wakil presiden yang bersaing, yaitu Anies Baswedan dengan Muhaimin Iskandar, Prabowo Subianto dengan Gibran Rakabuming Raka, serta Ganjar Pranowo dengan Mahfud MD[6][7]. Untuk memahami strategi komunikasi digital yang digunakan oleh setiap pasangan Capres dan Cawapres dalam konteks komunikasi dan isu-isu yang terkait, penting untuk mengeksplorasi peluang digitalisasi. Hal ini memungkinkan analisis terhadap pola komunikasi dan model topik yang diadopsi atau dikaitkan dengan masing-masing pasangan.

Menghadapi Pemilihan Presiden 2024, Prabowo Subianto mengadopsi strategi kampanye baru untuk menarik pemilih muda. Ia

berusaha menampilkan sisi yang lebih ramah dan menggemaskan melalui strategi "Gemoy". Dengan menunjukkan perilaku yang lebih dekat dan menyenangkan, Prabowo berharap memperluas basis pemilihnya[8]. Strategi kampanye Anies-Muhaimin fokus pada pendekatan langsung ke masyarakat melalui kampanye *door to door*, dialog dengan tokoh masyarakat, dan kunjungan ke daerah. Memanfaatkan media sosial secara intensif untuk menjangkau pemilih lebih luas, dengan konten menarik dan informatif. Ganjar-Mahfud mengutamakan dukungan generasi Z, dengan fokus pada isu lingkungan, kesehatan mental, pendidikan, lapangan kerja, serta seni dan budaya. Pemanfaatan Media sosial dan merencanakan creative hub untuk mendukung ide kreatif anak muda[9]. Dalam penelitian ini, metode web scraping digunakan untuk mengumpulkan data dari media online. Teknik ini memungkinkan ekstraksi otomatis informasi dari situs web dengan kata kunci tertentu. Data diambil dari Detik.com menggunakan kata kunci terkait Pilpres Indonesia 2024, seperti nama tim kampanye dan pasangan calon presiden[2][10].

Topic modeling adalah metode yang digunakan untuk mengidentifikasi topik-topik dalam serangkaian teks, bagian dari bidang Pengolahan Bahasa Alami (*Natural Language Processing*). Menganalisis teks menggunakan algoritma untuk menemukan pola-pola tersembunyi[11]. Pendekatan pemodelan yang digunakan adalah algoritma *Latent Dirichlet Allocation* (LDA). Algoritma LDA telah banyak digunakan dalam analisis teks, membantu mengenali pola, topik, dan relasi antar dokumen. Algoritma ini juga dapat diterapkan pada berbagai kasus seperti rekayasa perangkat

lunak[12][13]. Penggunaan *web scraping* dan *LDA* diharapkan dapat mengungkap topik-topik yang paling banyak dibahas serta pola komunikasi selama Pilpres 2024. Hasil penelitian ini diharapkan memberikan wawasan tentang strategi komunikasi digital yang digunakan oleh setiap pasangan Capres dan Cawapres, dengan menganalisis pola komunikasi dan model topik yang diadopsi atau dikaitkan dengan masing-masing pasangan.

2. TINJAUAN PUSTAKA

Analisis trending topik di media sosial telah banyak diterapkan untuk mengetahui tentang sentiment berbasis teks[14], juga tentang analisis persepsi publik[15] berbagai penelitian menggunakan algoritma *LDA* mengeksplorasi penerapan *LDA* untuk mengidentifikasi topik di grup *WhatsApp* Universitas Islam Majapahit, dengan tambahan dimensi waktu. Hasil terbaik menunjukkan precision 0.9294, recall 0.7900, dan f-measure 0.8541 [16]. Penerapan *LDA* untuk memodelkan topik faktor perceraian dari dokumen Pengadilan Agama Yogyakarta 2020. Hasilnya menunjukkan mayoritas kasus adalah cerai gugat (78%). Analisis *LDA* mengidentifikasi tiga topik utama, meninggalkan pasangan, masalah ekonomi, dan perlakuan terhadap salah satu pihak, dengan koherensi 0.29627 [17].

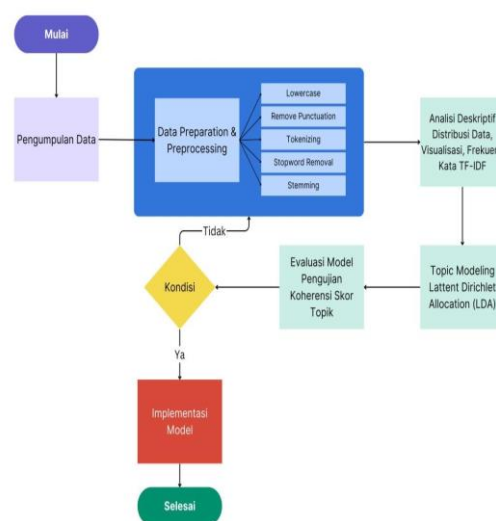
Penerapan *LDA* untuk mengidentifikasi topik dalam cuitan penyakit tropis di Indonesia, khususnya kusta, malaria, dan demam berdarah. Ada lima topik utama dengan koherensi 0.564049, termasuk penggunaan dana, isu Covid-19, pengobatan, perawatan, dan informasi tentang buta dan kusta[18]. Tiga penelitian berbeda menerapkan algoritma *LDA* dalam konteks yang berbeda. Pertama, *LDA* digunakan untuk menganalisis trending topik di media sosial, dengan fokus pada grup *WhatsApp* di Universitas Islam Majapahit. Hasilnya menunjukkan tingkat presisi, recall, dan f-measure yang tinggi. Kedua, *LDA* diterapkan untuk memodelkan topik faktor perceraian dari dokumen Pengadilan Agama Yogyakarta 2020, menemukan tiga topik utama yang dominan. Ketiga, *LDA* digunakan untuk mengidentifikasi

topik dalam cuitan tentang penyakit tropis di Indonesia, menyoroti lima topik utama yang sering dibicarakan, termasuk isu-isu seputar Covid-19. Selanjutnya, penelitian yang usulkan dengan judul Topik modeling pola komunikasi Pilpres 2024 menggunakan *web scraping* dan latent dirichlet allocation. Tujuan dari penelitian ini Menemukan topik dan pola komunikasi Pilpres 2024. Hasil penelitian menunjukkan bahwa model *LDA* memiliki tingkat koherensi 0,450 hal ini menunjukkan tingkat koherensi yang sedang, dengan topik-topik yang dominan berkisar pada dukungan terhadap kandidat, hasil survei, dan isu-isu politik terkait.

3. METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Penelitian ini bertujuan menganalisis model pengetahuan untuk memahami topik dan pola komunikasi yang digunakan oleh tim pemenangan dalam Pilpres 2024. Melalui pendekatan ini, penelitian mengidentifikasi topik dominan dan pola komunikasi yang digunakan dalam menyampaikan pesan politik kepada publik. Pemahaman model pengetahuan ini diharapkan memberikan wawasan mendalam tentang strategi dan taktik komunikasi politik yang efektif dalam konteks Pilpres.



Gambar 1. Tahapan Penelitian



Gambar 1 menunjukkan langkah-langkah detail dalam tahapan penelitian. Berikut adalah penjelasan lanjutan untuk setiap tahapan:

1. Pengumpulan Data

Tahap ini melibatkan proses mengumpulkan data dari berbagai sumber yang relevan, datasets yang digunakan pada penelitian ini bersumber dari situs media massa Detik.com dari tahun 2024 sampai dengan tahun 2023[4].

2. Preparation dan preprocessing of data

Setelah data terkumpul, langkah selanjutnya adalah membersihkan dan mempersiapkan data untuk analisis, tahapan ini melibatkan lima tahapan, yaitu lowercasing, remove punctuation, tokenizing, stopword removal, dan stemming.

3. Analisis Deskriptif dan Visualisasi Data

Tahap ini mencakup analisis statistik deskriptif untuk memahami karakteristik dasar dari data, seperti jumlah berita, kata pendek dan sebaran frekuensi kata Selain itu, visualisasi data seperti barchart, wordcloud dan pyLDAvis juga digunakan untuk memvisualisasikan distribusi data

4. Frekuensi kata *TF-IDF*

Tahap selanjutnya yaitu pembobotan *TF-IDF* yang digunakan untuk mengevaluasi seberapa pentingnya sebuah kata di dalam sebuah dokumen. *Term Frequency (TF)* yaitu semakin tinggi frekuensi kemunculan term pada sebuah dokumen maka akan semakin tinggi juga nilai bobot untuk *term* itu sendiri. Sementara itu, proses *Inverse Document Frequency (IDF)* merupakan kebalikan dari proses *TF*. Pada

IDF, semakin tinggi frekuensi kemunculan term maka nilai bobot *term* itu sendiri akan semakin kecil. Berikut persamaan *TF-IDF* yang digunakan pada penelitian ini. Berikut rumus untuk mencari nilai *TF-IDF* [21] :

$$W_{dt} = tf_d \times idf_t \quad (1)$$

$$idf_t = \log \frac{N}{df_t}$$

Keterangan:

W_{dt} = bobot dokumen ke-d terhadap kata ke-t

tf_d = banyaknya kata yang dicari pada sebuah dokumen

N = Jumlah dokumen secara keseluruhan

df_t = Jumlah dokumen yang mengandung kata t

d = Dokumen ke-d

t = Kata ke-t

5. Pemodelan topik Latent Dirichlet Allocation (LDA)

Proses pemodelan dengan menggunakan *Latent Dirichlet Allocation* melibatkan beberapa tahapan yang penting. Pertama, kita harus menentukan jumlah topik yang ingin diidentifikasi dalam koleksi dokumen tersebut. Selanjutnya, LDA memperkirakan distribusi topik dalam setiap dokumen, yang disebut theta, dan distribusi kata dalam setiap topik, yang disebut beta. Langkah selanjutnya adalah menentukan topik untuk setiap kata dalam setiap dokumen. Ini dilakukan dengan memperhitungkan probabilitas bahwa sebuah kata diberikan topik tertentu, yang dinyatakan dengan rumus LDA sebagai berikut [19]:

$$\rho(w, z, \theta, \alpha, \beta) = \quad (2)$$

$$\prod_{d=1}^D \rho(\theta_d | \alpha) \prod_{n=1}^{Nd} \rho(w_{d,n} | \theta_d, \beta_{1:K})$$

Keterangan :

(D) adalah jumlah dokumen dalam koleksi.

(Nd) adalah jumlah kata dalam dokumen (d).

($w_{d,n}$) adalah kata ke-(n) dalam dokumen (d).

(z, d, n) adalah topik yang terkait dengan kata ($w_{d,n}$).

(θ) adalah distribusi topik dalam dokumen (d).

(α) dan (β) adalah parameter model.

Rumus tersebut menggambarkan probabilitas keseluruhan dari semua variabel dalam model LDA, termasuk distribusi topik dalam setiap dokumen ($\rho(\theta_d | \alpha)$) dan distribusi kata dalam setiap topik ($\rho(w_{d,n} | \theta_d, \beta_{1:K})$). Penerapan LDA tidak hanya melibatkan perhitungan matematis yang kompleks, tetapi juga algoritma pembelajaran yang tepat untuk

mengestimasi parameter-parameter yang terlibat. Beberapa teknik pembelajaran yang umum digunakan termasuk variational inference dan Gibbs sampling. Dengan menggunakan pendekatan ini, LDA dapat menghasilkan estimasi yang cukup akurat tentang distribusi topik dalam dokumen-dokumen yang diberikan. Dalam praktiknya, penerapan LDA membawa manfaat yang signifikan dalam analisis teks. Ini tidak hanya memungkinkan penemuan pola dan tematik yang tersembunyi dalam koleksi dokumen, tetapi juga dapat digunakan untuk tujuan seperti pengelompokan dokumen, analisis sentimen, klasifikasi teks, dan masih banyak lagi. Dengan menyediakan cara yang sistematis untuk mengidentifikasi topik dalam dokumen, LDA telah menjadi salah satu alat utama dalam penambangan teks dan pemrosesan bahasa alami. Penerapan LDA tidak hanya melibatkan perhitungan matematis yang kompleks, tetapi juga algoritma pembelajaran yang tepat untuk mengestimasi parameter-parameter yang terlibat. Beberapa teknik pembelajaran yang umum digunakan termasuk variational inference dan Gibbs sampling. Dengan menggunakan pendekatan ini, LDA dapat menghasilkan estimasi yang cukup akurat tentang distribusi topik dalam dokumen-dokumen yang diberikan.

6. Evaluasi Model

Pada tahap ini peneliti melakukan pengujian untuk mencari dan menghitung nilai koherensi dalam pemrosesan teks untuk memetakan sejauh mana topik yang dihasilkan oleh (3) model yang dianggap dapat koheren atau saling berkaitan. Nilai koherensi dalam pemodelan topik digunakan untuk mengukur seberapa mudah topik tersebut ditafsirkan oleh manusia. Nilai koherensi digunakan untuk membantu menentukan berapa banyak topik yang sesuai dari dokumen. Semakin tinggi nilai koherensi, semakin baik hasil interpretasi pemodelan topik. Adapun rumus untuk menghitung nilai koherensi adalah sebagai berikut:

$$\text{Coherence (V)} = \frac{\sum_{(v_i, v_j) \in V} \text{Score}(v_i, v_j, \epsilon)}{|V|} \quad (3)$$

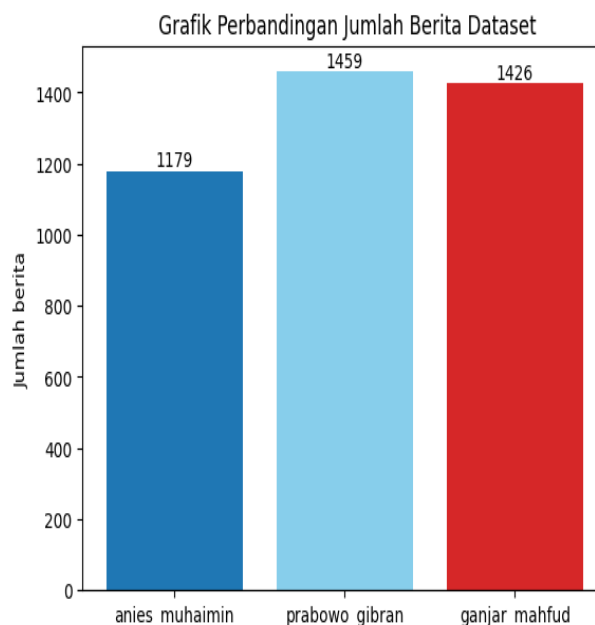
Keterangan :

V=kumpulan kata untuk menjelaskan sebuah topik. ϵ =adalah faktor smoothing untuk mengembalikan nilai ke bilangan riil [20].

4. HASIL DAN PEMBAHASAN

4.1 Analisa Deskriptif

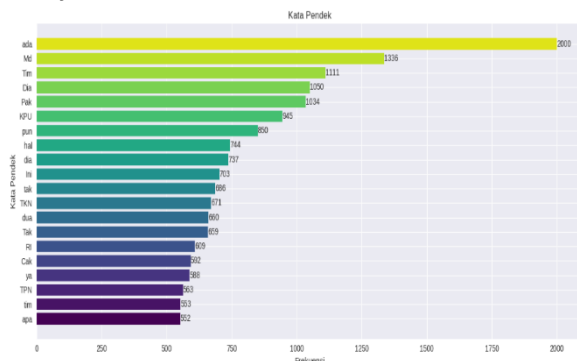
Analisis deskriptif dilakukan untuk mendalami data, memahami bentuk, struktur, dan karakteristik dalam dataset. Total data yang dianalisis berjumlah 4.064 dari tahun 2023 hingga 2024, dengan rincian: dataset Anies-Muhaimin sebanyak 1.179, dataset Prabowo-Gibran sebanyak 1.459, dan dataset Ganjar-Mahfud sebanyak 1.426. Gambar 4.1 menunjukkan perbandingan jumlah ketiga dataset di atas.



Gambar 2. Perbandingan Jumlah Dataset

Grafik ini menunjukkan 20 kata pendek yang paling sering digunakan, diurutkan dari frekuensi tertinggi hingga terendah. Dari grafik tersebut, dapat diinterpretasikan bahwa kata "ada" muncul paling sering dalam dataset Pilpres 2024, dengan frekuensi 2000 kali. Kata "MD" muncul

untuk menilai pentingnya kata berdasarkan frekuensi dan keunikan di antara dokumen.



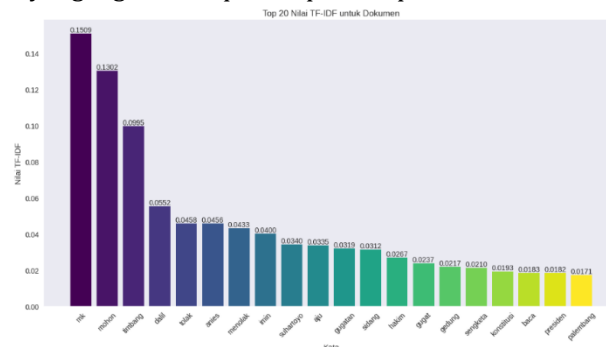
Gambar 3. Visualisasi Kata Pendek

[illegible]

Gambar 5. Visualisasi Wordcloud

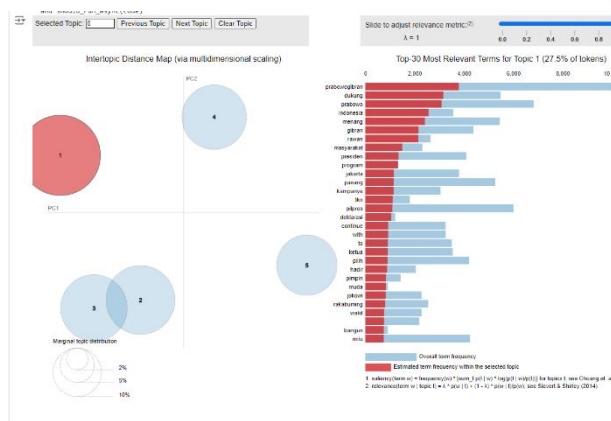
Gambar 4. Daftar *Rank* Dan *Term* Kata

Hasil analisis menggunakan metode TF-IDF memberikan pemahaman yang mendalam tentang pentingnya setiap kata dalam teks yang dianalisis. Gambar 4.4 menjelaskan 20 jenis kata yang memiliki bobot tinggi dalam analisis menggunakan TF-IDF. Hasil visualisasi TF-IDF memberikan pemahaman yang jelas tentang urutan kata-kata kunci berdasarkan bobot tertinggi, yang mencakup kata-kata seperti "mk", "mohon", "timbang", "dalil", "tolak", "anies", "menolak", "imin", "suhartoyo", "aju", "sidang", "hakim", "gugat", "gedung", "sengketa", "konstitusi", "baca", "presiden" dan "palembang". Dengan melihat kata-kata tersebut, topik atau tema yang paling dominan dalam dokumen dapat diidentifikasi dengan cepat, sehingga memberikan wawasan yang berharga dalam memahami esensi dari teks dan memungkinkan pembaca (*stakeholder*) menangkap inti dari topik yang ingin disampaikan pada Pilpres 2024.



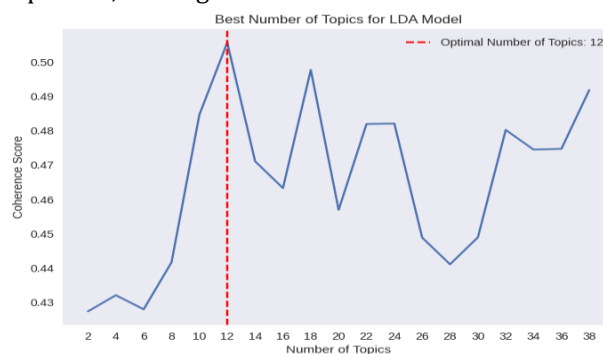
Gambar 6. Visualisasi nilai TF-IDF

Setelah TF-IDF, langkah berikutnya adalah pemodelan topik dengan LDA untuk memahami struktur topik teks lebih dalam. LDA membantu mengidentifikasi pola tersembunyi dan mengelompokkan kata-kata ke dalam topik yang terkait, memberikan gambaran mendalam tentang isi teks. Hasilnya menunjukkan lima topik utama. Kesimpulan menyoroti pentingnya umpan balik pengguna untuk meningkatkan pengalaman pengguna.



Gambar 7. Visualisasi Topik Model

Dengan koherensi 0.405, model LDA menunjukkan tingkat konsistensi sedang, menandakan hubungan yang signifikan antar kata dalam setiap topik. Perbaikan dapat dilakukan melalui penyesuaian parameter atau preprocessing data. Tingkat konsistensi yang tinggi memastikan topik terstruktur dan mudah dipahami, meningkatkan wawasan dari data.



Model LDA menunjukkan konsistensi sedang dengan koherensi 0.405, menandakan hubungan yang signifikan antar kata dalam setiap topik. Terdapat ruang untuk peningkatan model melalui penyesuaian parameter atau perbaikan preprocessing data. Tingkat koherensi yang lebih tinggi diharapkan memperbaiki pemahaman topik dan wawasan dari data. Diperlukan penelitian lanjutan dan pengembangan model yang lebih cermat. Dalam hasil pemodelan LDA, topik utama dalam Pilpres 2024 mencakup dukungan terhadap kandidat, hasil survei, dan isu-isu terkait. Rekomendasi untuk penelitian selanjutnya termasuk penggunaan data yang lebih luas dan pengumpulan informasi dari media sosial. Dari koherensi 0.405, model LDA menunjukkan konsistensi sedang. Perbaikan model bisa dilakukan melalui penyesuaian parameter atau preprocessing data. Ini diharapkan akan meningkatkan pemahaman topik dan wawasan dari data yang dianalisis. Perlu penelitian lanjutan dan pengembangan model yang lebih cermat untuk analisis yang lebih akurat. Hasil pemodelan LDA menunjukkan topik utama dalam Pilpres 2024 untuk setiap pasangan

calon. Rekomendasi untuk penelitian selanjutnya meliputi penggunaan data yang lebih luas dan perbaikan preprocessing data. Integrasi data media sosial seperti Twitter, Instagram, atau Youtube dianjurkan untuk pemahaman yang lebih komprehensif.

Tabel 1. Topic Pilpres

Topic Number	Hasil Word	Interpretasi Topik
Topic 0 : Pemilihan Presiden (Pilpres) dan Keputusan Mahkamah Konstitusi (MK)	0.014*"pilpres" + 0.014*"imin" + 0.013*"mk" + 0.012*"anies" + 0.010*"cak" + 0.010*"amin" + 0.009*"presiden" + 0.009*"putus" + 0.009*"kpu" + 0.009*"jakarta"),	Topik ini tampaknya berkaitan dengan pemilihan presiden (pilpres) dan keputusan Mahkamah Konstitusi (MK) terkait dengan hal tersebut. Kata kunci seperti "pilpres", "MK", "anies", "amin", dan "kpu" menunjukkan topik ini berfokus pada proses dan hasil pemilihan presiden.
Topic 1 : Perhitungan Suara dan Hasil Pemilihan	0.050*"suara" + 0.021*"ganjarmahfud" + 0.020*"prabowogibran" + 0.015*"hasil" + 0.015*"kpu" + 0.011*"tps" + 0.011*"hitung" + 0.011*"pilpres" + 0.010*"pasang" + 0.010*"nomor")],	Topik ini berkaitan dengan perhitungan suara dan hasil pemilihan, dengan kata kunci seperti "suara", "ganjarmahfud", "prabowogibran", "hasil", dan "kpu". Hal ini menunjukkan topik ini berfokus pada analisis hasil suara pemilihan dan proses perhitungannya.
Topic 2 : Survei Opini dan Dukungan Terhadap Pasangan Calon	0.026*"prabowogibran" + 0.025*"survei" + 0.013*"prabowo" + 0.013*"pilih" + 0.012*"dukung" + 0.011*"pasang" + 0.010*"menang" + 0.010*"pilpres" +	Topik ini berkaitan dengan survei opini dan dukungan terhadap pasangan calon, dengan kata kunci seperti

	0.009*"hasil" + 0.009*"jatim")],	"survei", "prabowogibran", "pilih", dan "dukung". Topik ini mungkin membahas hasil survei opini terkini dan dukungan publik terhadap pasangan calon tertentu.
Topic 3 : Dukungan terhadap Prabowo-Gibran dan Isu Kekawatiran	0.020*"prabowogibran" + 0.016*"dukung" + 0.016*"prabowo" + 0.013*"indonesia" + 0.012*"menang" + 0.011*"gibran" + 0.011*"rawan" + 0.008*"masyarakat" + 0.007*"presiden" + 0.007*"program")],	Topik ini berkaitan dengan dukungan terhadap pasangan Prabowo-Gibran serta isu-isu kekhawatiran terkait, dengan kata kunci seperti "prabowogibran", "dukung", "indonesia", dan "rawan". Isu-isu tersebut mungkin meliputi dukungan politik, potensi risiko, dan isu-isu keamanan.
Topic 4 : Kegiatan Kampanye dan Partai Politik	0.017*"ganjar" + 0.015*"ganjarmahfud" + 0.013*"mahfud" + 0.011*"kampanye" + 0.009*"debat" + 0.007*"anies" + 0.007*"md" + 0.007*"ketua" + 0.007*"tpn" + 0.006*"pdip")]	Topik ini tampaknya berkaitan dengan kegiatan kampanye dan partai politik, dengan kata kunci seperti "ganjar", "ganjarmahfud", "mahfud", "kampanye", dan "debat". Topik ini mungkin membahas kegiatan kampanye dari para kandidat dan peran partai politik dalam proses pemilihan.



5.1 KESIMPULAN

Penelitian ini menyoroti pentingnya komunikasi politik dalam Pemilihan Umum Presiden (Pilpres) di Indonesia, khususnya dalam konteks Pilpres 2024 yang diikuti oleh tiga pasangan calon: Anies Baswedan dengan Muhaimin Iskandar, Prabowo Subianto dengan Gibran Rakabuming Raka, serta Ganjar Pranowo dengan Mahfud MD. Setiap pasangan calon mengadopsi strategi komunikasi digital yang berbeda. Prabowo Subianto, misalnya, menggunakan strategi "Gemoy" untuk menarik pemilih muda dengan menampilkan sisi yang lebih ramah. Anies-Muhaimin fokus pada pendekatan langsung melalui kampanye *door to door* dan dialog dengan masyarakat, serta memanfaatkan media sosial untuk menjangkau lebih banyak pemilih. Sementara itu, Ganjar-Mahfud menargetkan generasi Z dengan menekankan isu-isu lingkungan, kesehatan mental, pendidikan, lapangan kerja, serta seni dan budaya, dan merencanakan creative hub untuk mendukung ide kreatif anak muda. Metode *web scraping* digunakan untuk mengumpulkan data dari media online, khususnya dari Detik.com, untuk menganalisis pola komunikasi dan model topik yang terkait dengan Pilpres 2024. *Topic modeling* dengan algoritma *Latent Dirichlet Allocation (LDA)* digunakan untuk mengidentifikasi topik-topik utama dalam teks yang dianalisis. Hasilnya menunjukkan konsistensi sedang dengan koherensi 0.405, yang menunjukkan hubungan yang signifikan antar kata dalam setiap topik. Hasil analisis menunjukkan lima topik utama terkait Pilpres 2024 di Indonesia. Pertama, pemilihan presiden dan keputusan Mahkamah Konstitusi yang berfokus pada proses dan hasil pemilihan. Kedua, perhitungan suara dan hasil pemilihan, termasuk analisis suara yang diterima oleh pasangan calon. Ketiga, survei opini dan dukungan terhadap pasangan calon, mencakup hasil survei dan preferensi pemilih. Keempat, dukungan dan kekhawatiran terkait pasangan Prabowo-Gibran, meliputi isu-isu politik dan keamanan. Kelima,

kegiatan kampanye dan peran partai politik, yang menyoroti aktivitas kampanye para kandidat dan kontribusi partai politik dalam proses pemilihan.

5.2 Saran

Penelitian ini memberikan wawasan tentang strategi komunikasi digital yang digunakan oleh setiap pasangan calon presiden dan wakil presiden. Namun, untuk meningkatkan pemahaman topik dan wawasan dari data yang dianalisis, diperlukan penelitian lanjutan dengan penyesuaian parameter atau perbaikan preprocessing data. Integrasi data dari media sosial seperti Twitter, Instagram, dan YouTube juga dianjurkan untuk memberikan pemahaman yang lebih komprehensif.

DAFTAR PUSTAKA:

- [1] Topan Yuniarto, "Pemilihan Umum 2024 dalam Angka – Kompaspedia." KOMPASPEDIA, 2024.
- [2] T. Komunikasi, P. Di, and M. Sosial, "Jurnal Lanskap Politik Jurnal Lanskap Politik," vol. 1, 2023.
- [3] M. F. I Islami, "OPINI_ Pengaruh Media Sosial Terhadap Pemilu 2024 - Prodi Komunikasi Penyiaran Islam." kpi.iainpare.ac.id, 2023.
- [4] V. N. April *et al.*, "Media Sosial sebagai Agen Transformasi Politik: Analisis Pengaruh terhadap Proses Komunikasi Politik Dinamika Komunikasi Politik dalam Era Media Sosial: Peran Platform Digital dalam Pembentukan Opini Publik," mengeksplorasi peran penting yang dimainkan oleh platform diskursus politik, menghilangkan hambatan tradisional antara pemimpin dan pemilih. Oleh politik dalam era media sosial dan menggali peran kritis yang dimainkan oleh platform digital," vol. 2, no. 2, 2024.
- [5] A. S. Cahyono, "Anang Sugeng Cahyono, Pengaruh Media Sosial Terhadap Perubahan Sosial Masyarakat di Indonesia," pp. 140–157.
- [6] B. N. INDONESIA, "Pemilu 2024: Tiga pasangan capres-cawapres akan bertarung dalam Pilpres, apa saja visi dan misi mereka?," *BBC NEWS INDONESIA*, 2023.
- [7] Andrean W. Finaka, "Profil Calon Presiden dan Wakil Presiden di Pemilu 2024 – Indonesia Baik." indonesiabaik.id, 2024.
- [8] K. Y. R. WILLY MEDI CHRISTIAN



- NABABAN, "Strategi "Gemoy" Prabowo di Tengah Gempuran Sentimen Negatif," *kompas*, 2023.
- [9] A. P. Surya, "Strategi Kampanye Tiap Calon pada Pemilu 2024 Serta Propaganda yang Terjadi pada Pasangan AMIN Konten ini telah tayang di Kompasiana.com dengan judul 'Strategi Kampanye Tiap Calon pada Pemilu 2024 Serta Propaganda yang Terjadi pada Pasangan AMIN,'" *kompasiana*, 2023.
- [10] S. Chohan, A. Nugroho, A. M. Bayu Aji, and W. Gata, "Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naïve Bayes dan Synthetic Minority Over Sampling Technique," *Paradig. Inform. dan Komput.*, vol. 22, no. 2, pp. 139–144, 2020.
- [11] K. R. A. P. Santoso, A. Husna, N. W. Putri, and N. A. Rakhmawati, "Analisis Topik Tagar Covidindonesia pada Instagram Menggunakan Latent Dirichlet Allocation," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 7, no. 1, pp. 1–9, 2022.
- [12] A. R. D. Astuti and N. Cahyono, "Analisis Topic Modelling Persepsi Pengguna Internet Menggunakan Metode Latent Dirichlet Allocation," *Indones. J. Comput. Sci.*, vol. 12, no. 1, 2023.
- [13] W. W. Astuti, "Pelatihan Pembuatan Soal Menggunakan Google Formulir Di SDN Batangkaluku," *J. Train. Community Serv. Adptersi*, vol. 2, no. 1, pp. 27–32, 2022.
- [14] M. Nitami and H. Februriyanti, "Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma Naïve Bayes," *J. Manaj. dan Sist. Inf.*, vol. 7, no. 1, pp. 20–29, 2022.
- [15] I. W. D. Pancane and I. W. Suriana, "Klasifikasi Komentar Publik Terhadap Kebijakan Pemerintah Pada Facebook Frontpage Kompas Menggunakan Clustering K-Means Furthest First," *J. Manaj. Inform. dan Sist. Inf.*, vol. 3, no. 2, pp. 166–173, 2020.
- [16] A. Syaifuddin, R. A. Harianto, and J. Santoso, "Analisis Trending Topik untuk Percakapan Media Sosial dengan Menggunakan Topic Modelling Berbasis Algoritme LDA," *INSYST J. Intell. Syst. Comput.*, vol. 2, no. 1, pp. 12–19, 2020.
- [17] E. P. Putri, "Mplementasi Latent Dirichlet Allocation (Lda) untuk Pemodelan Topik Faktor Perceraian (Studi Kasus: Dokumen Putusan Perceraian di Pengadilan Agama Yogyakarta Tahun 2020)." Universitas Islam Indonesia, 2022.
- [18] D. Ridhwanullah, "Pemodelan Topik pada Cuitan tentang Penyakit Tropis di Indonesia dengan Metode Latent Dirichlet Allocation," 2022.
- [19] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, no. Jan, pp. 993–1022, 2003.
- [20] K. Stevens, P. Kegelmeyer, D. Andrzejewski, and D. Buttler, "Exploring topic coherence over many models and many topics," in *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, 2012, pp. 952–961.