

ANALISIS PERSEPSI PUBLIK TERHADAP RESHUFFLE MENTERI KEUANGAN 2025 BERDASARKAN TOPIC MODELING LDA PADA KOMENTAR YOUTUBE

Nayoan fajar ade cendra, Nur Kholis Wakhid, Harun Al Rosyid

Pendidikan Teknologi Informasi, Fakultas Teknik, Universitas Negeri Surabaya
Jl. Ketintang Wiyata Gedung A10, Ketintang, Gayungan, Surabaya, Jawa Timur 60231
nayoan.22076@mhs.unesa.ac.id, Nayoan Fajar Ade Chendra

ABSTRAK

Perkembangan media sosial, khususnya YouTube, menjadikannya ruang publik yang penting dalam pembentukan dan penyebaran opini masyarakat terhadap isu-isu strategis nasional. Salah satu isu yang mendapat perhatian luas pada tahun 2025 adalah reshuffle Menteri Keuangan, yang memicu beragam respons publik dalam kolom komentar YouTube. Permasalahan utama dalam penelitian ini adalah bagaimana mengidentifikasi dan memetakan persepsi publik yang kompleks dari data komentar berjumlah besar secara sistematis dan objektif. Penelitian ini bertujuan untuk menganalisis persepsi publik terhadap reshuffle Menteri Keuangan 2025 menggunakan pendekatan topic modeling. Metode penelitian meliputi pengumpulan 51.550 komentar dari 437 video YouTube menggunakan YouTube Data API v3, dilanjutkan dengan tahapan preprocessing berbasis Natural Language Processing dan pemodelan topik menggunakan algoritma Latent Dirichlet Allocation (LDA). Hasil penelitian menunjukkan terbentuknya tiga topik utama, yaitu harapan publik terhadap kepemimpinan Menteri Keuangan (35,2%), kritik terhadap tata kelola keuangan dan isu korupsi (28,4%), serta dinamika politik dan reshuffle kabinet (36,4%). Evaluasi model menghasilkan nilai coherence rata-rata 0,5586 dan topic diversity sebesar 0,833, yang menunjukkan kualitas topik yang baik. Temuan ini menegaskan bahwa komentar YouTube dapat digunakan sebagai sumber data representatif untuk memahami persepsi publik terhadap isu politik strategis.

Kata kunci : *topic modeling, LDA, persepsi publik, YouTube comments, natural language processing*

1. PENDAHULUAN

Perkembangan teknologi informasi dalam dua dekade terakhir telah mengubah cara masyarakat mengakses, mendiskusikan, dan merespons isu-isu publik. Media sosial kini tidak hanya berfungsi sebagai sarana hiburan, tetapi juga sebagai ruang diskusi dan pertukaran informasi yang mempengaruhi pembentukan opini publik [1]. Di antara berbagai platform digital yang ada, YouTube merupakan salah satu media terbesar dengan jutaan pengguna aktif setiap harinya. Fitur kolom komentar pada YouTube memungkinkan pengguna untuk menyampaikan pandangan, kritik, dukungan, maupun reaksi spontan terhadap konten yang mereka konsumsi [2]. Kondisi ini menjadikan komentar YouTube sebagai sumber data yang sangat potensial untuk mengkaji persepsi publik secara empiris dan berbasis data.

Pada tahun 2025, wacana reshuffle Menteri Keuangan menjadi salah satu topik yang menimbulkan perhatian luas di tengah masyarakat. Pergantian pada jabatan strategis ini tidak hanya memiliki implikasi terhadap kebijakan ekonomi nasional, tetapi juga mencerminkan dinamika politik dalam pemerintahan [3]. Oleh karena itu, publik menunjukkan beragam reaksi mulai dari rasa penasaran, kekhawatiran, dukungan, hingga kritik yang intens. Diskusi tersebut banyak berlangsung di ruang digital, khususnya YouTube, melalui konten berita, analisis politik, dan opini publik yang membahas isu reshuffle tersebut [4]. Banyaknya komentar yang muncul mencerminkan kompleksitas persepsi publik, namun belum seluruhnya dianalisis secara sistematis, sehingga

menimbulkan tantangan dalam memahami pola opini masyarakat dari data teks berskala besar [5].

Dalam menghadapi permasalahan tersebut, diperlukan pendekatan analitis yang mampu mengolah dan menafsirkan data teks secara efisien dan objektif. Analisis teks dan topic modeling menawarkan solusi komputasional untuk mengeksplorasi pola komunikasi publik secara sistematis. Topic modeling, khususnya metode Latent Dirichlet Allocation (LDA), digunakan untuk mengidentifikasi topik-topik utama yang muncul dari kumpulan komentar tanpa harus membaca setiap komentar secara manual [6]. Selain itu, pendekatan Natural Language Processing (NLP) memungkinkan proses pembersihan data, identifikasi kata kunci, serta pemetaan konteks bahasa yang digunakan dalam diskusi publik [7].

Berdasarkan latar belakang dan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis persepsi publik terhadap reshuffle Menteri Keuangan melalui komentar pada video YouTube yang membahas peristiwa reshuffle menteri pada 8 September 2025. Penelitian dilakukan melalui tahapan pengumpulan data (scraping komentar), preprocessing data teks, penerapan topic modeling menggunakan LDA, serta interpretasi hasil analisis untuk mengidentifikasi topik dominan dan kecenderungan opini publik.

Hasil penelitian ini diharapkan mampu memberikan gambaran yang komprehensif mengenai struktur diskusi dan persepsi masyarakat di ruang digital, sekaligus berkontribusi pada pengembangan kajian analisis media sosial dalam konteks isu politik di Indonesia. Selain itu, temuan penelitian ini

diharapkan dapat menjadi bahan pertimbangan bagi pemangku kebijakan dalam merumuskan kebijakan publik yang lebih responsif terhadap aspirasi dan opini masyarakat yang berkembang di media sosial.

2. TINJAUAN PUSTAKA

2.1. Media Sosial dan Opini Publik

Media sosial telah merubah cara masyarakat dalam mengakses dan mendiskusikan isu-isu politik. YouTube menjadi platform signifikan dengan lebih dari 2 miliar pengguna aktif, berfungsi tidak hanya sebagai sarana hiburan tetapi juga ruang deliberasi informal di mana publik dapat menyuarakan pandangan, kritik, dan aspirasi mereka terhadap peristiwa-peristiwa strategis[8],[9]. Fitur kolom komentar pada YouTube memungkinkan pengguna untuk memberikan respons spontan dan reflektif terhadap konten, menjadikannya sumber data yang sangat potensial untuk memahami persepsi masyarakat tentang kebijakan pemerintah[10]

Diskusi public di media sosial sering melibatkan dimensi aspirasional (harapan terhadap kepemimpinan), kritis (evaluasi kinerja), dan spekulatif (analisis motif politikus)[11],[12]. Penelitian menunjukkan bahwa Ketika menghadapi pergantian kepemimpinan strategis, public mengekspresikan harapan idealistis sambil melakukan analisis struktural terhadap masalah kebijakan, mencerminkan meningkatnya literasi ekonomi dan politikus di era digital[13].

2.2. Analisis Teks dan Natural Language Processing

Analisis teks berbasis *Natural Language Processing* (NLP) memungkinkan ekstraksi informasi bermakna dari data teks tidak terstruktur dalam volume besar [14]. Pipeline NLP mencakup preprocessing, tokenisasi, normalisasi, stopword removal, dan stemming, dengan setiap tahapan dirancang untuk membersihkan teks dari elemen tidak relevan dan menormalisasi bentuk kata agar siap untuk analisis komputasional [15],[16].

Tantangan khusus dalam NLP untuk bahasa Indonesia mencakup kompleksitas morfologi, keragaman dialek, dan code-switching dalam komunikasi digital [17]. Penggunaan stemmer Indonesia seperti Sastrawi dan customized preprocessing pipeline yang disesuaikan dengan karakteristik data spesifikasi terbukti meningkatkan akurasi analisis sentimen dan topic modeling pada data media sosial berbahasa Indonesia [18],[19]

2.3. Topic Modeling dan Latent Dirichlet Allocation

Topic modelling adalah Teknik unsupervised learning untuk menemukan struktur tematik tersembunyi dalam Kumpulan dokumen besar tanpa memerlukan label manual [20]. *Latent Dirichlet Allocation* (LDA) merupakan algoritma probabilistik paling populer yang memodelkan setiap dokumen

sebagai distribusi topik dan setiap topik sebagai distribusi kata [21]. LDA menggunakan hyperparameter alpha (*document-topic density*) dan beta (*topic-word density*) yang memerlukan evaluasi iterative untuk menghasilkan hasil optimal [22].

Meskipun alternatif Berbasis neural networks seperti BERTopic dan Top2Vec telah dikembangkan, LDA tetap menjadi pilihan robust untuk dataset komentar YouTube dengan ukuran sedang karena interpretabilitas tinggi, computational efficiency, dan track record kesuksesan dalam penelitian serupa [23],[24],[25]. Pemilihan LDA dalam penelitian ini didasarkan pada pertimbangan praktis dan validitas metodologis untuk konteks data media sosial Indonesia

2.4. Evaluasi Topic Modeling

Evaluasi kualitas topik dilakukan menggunakan metrik coherence score (mengukur keterkaitan semantik antar kata dalam topik), dan topic diversity (mengukur keberagaman kata antar topik) [26],[27]. Coherence score tinggi mengindikasikan topik yang kohesif dan terpisah; namun, evaluasi manual dan validasi kualitatif tetap esensial karena metrik otomatis memiliki keterbatasan dalam menangkap aspek pragmatis dan kontekstual [28],[29].

Penelitian menunjukkan bahwa kombinasi metrik otomatis dan human interpretation memberikan hasil evaluasi yang paling komprehensif [30]. Dalam penelitian ini, validasi dilakukan dengan mengintegrasikan coherence score, topic diversity, dan interpretasi kualitatif untuk memastikan bahwa topik-topik yang dihasilkan tidak hanya memiliki kualitas statistik baik tetapi juga bermakna dan relevan dengan konteks diskusi public tentang reshuffle Menteri Keuangan 2025.

2.5. Peneletian Terdahulu

Penelitian mengenai analisis opini publik berbasis media sosial dengan pendekatan topic modeling telah banyak dilakukan. Studi berjudul “Topic Modelling Latent Dirichlet Allocation untuk Klasifikasi Komentar pada Layanan Streaming Platform” [31] memanfaatkan metode Latent Dirichlet Allocation (LDA) untuk mengidentifikasi topik dominan pada komentar pengguna media sosial Twitter. Penelitian tersebut menunjukkan bahwa LDA mampu merepresentasikan struktur topik secara efektif serta memiliki tingkat interpretabilitas yang baik. Namun demikian, hasil klasifikasi sentimen dipengaruhi oleh ketidakseimbangan distribusi data komentar positif dan negatif.

Penelitian lain berjudul “Penerapan Pemodelan Topik Komentar melalui Media Sosial Twitter Menggunakan Latent Dirichlet Allocation (Studi Kasus: Pemerintah Kota Malang)” [32] menerapkan LDA untuk menganalisis komentar masyarakat terhadap isu pemerintahan daerah. Penelitian ini mengevaluasi pengaruh penggunaan stemming dalam proses preprocessing terhadap kualitas model topik.

Hasil penelitian menunjukkan bahwa model LDA tanpa stemming menghasilkan nilai perplexity yang lebih rendah dan coherence score yang lebih tinggi, sehingga topik yang dihasilkan lebih koheren dan mudah diinterpretasikan.

Selain itu, penelitian mengenai analisis topik pada opini publik berbasis media sosial [33] menunjukkan bahwa LDA efektif digunakan untuk mengekstraksi tema-tema kebijakan publik dari data teks tidak terstruktur dalam jumlah besar. Penelitian lain yang mengombinasikan analisis sentimen dan topic modeling pada platform digital [34] juga menunjukkan bahwa pendekatan ini mampu memberikan gambaran komprehensif mengenai persepsi dan kecenderungan opini masyarakat terhadap suatu isu atau layanan tertentu.

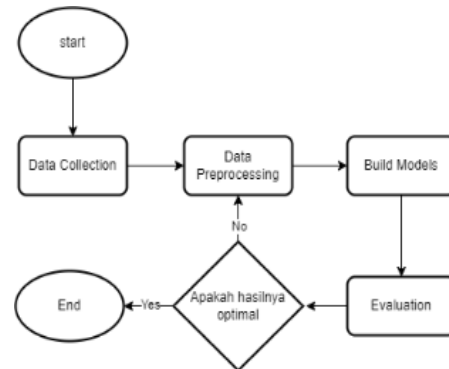
Berdasarkan penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa metode LDA terbukti efektif dalam menganalisis opini publik pada media sosial, khususnya pada platform berbasis teks seperti Twitter. Namun, sebagian besar penelitian masih berfokus pada isu layanan digital dan pemerintahan daerah. Oleh karena itu, penelitian ini memiliki kontribusi kebaruan dengan mengkaji persepsi publik terhadap reshuffle Menteri Keuangan tahun 2025 melalui komentar YouTube, yang memiliki karakteristik diskusi, panjang teks, serta dinamika interaksi yang berbeda. Penelitian ini diharapkan dapat memperluas kajian analisis opini publik berbasis media sosial dalam konteks isu politik nasional di Indonesia.

3. METODE PENELITIAN

3.1. Data Collection (Pengumpulan Data)

Pengumpulan data dilakukan menggunakan YouTube Data API v3 [35]. Pengambilan data ini

berfokus pada video yang membahas isu reshuffle Menteri Keuangan dalam rentang waktu mulai 8 September 2025 hingga dua minggu setelahnya. Melalui API ini, peneliti mengekstraksi komentar pengguna beserta metadata terkait, termasuk ID komentar, ID video, timestamp unggahan, dan jumlah likes, yang berfungsi sebagai variabel pendukung untuk analisis lebih lanjut [36].



Gambar 1. Alur Metodologi Penelitian

Untuk menjaga kualitas dan konsistensi data, dilakukan proses penyaringan awal dengan menghapus komentar yang bersifat duplikat, kosong, spam, atau hanya berisi emotikon [37]. Komentar yang lolos penyaringan kemudian diorganisasi dan disimpan dalam format CSV, sehingga memudahkan integrasi ke tahap preprocessing dan analisis lanjutan. Dari proses pengumpulan ini, diperoleh 437 video dan 51.550 komentar yang siap dianalisis untuk mendukung penelitian.

Kolom: ['video_id', 'judul', 'author', 'text', 'like_count', 'published_at']

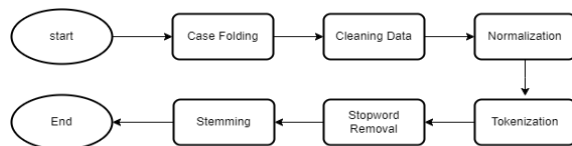
Jumlah data awal: 51550

Tabel 1. Sample Data YouTube Comments

video_id	judul	author	text	like_count	published_at
e0KFDSWcEOY	●LIVE - Menteri Keuangan Purbaya Umumkan Kondi...	@eryadimeilinton508	Pak. Buat kan pak. rekening pembayaran hutang n...	1	2025-09-22 09:22:41
e0KFDSWcEOY	●LIVE - Menteri Keuangan Purbaya Umumkan Kondi...	@aman-q1v8r	Assalamu'alaikum, ciamis hadir menyimak	1	2025-09-22 09:08:08
LJIoC4QAP9A	MENAKAR RESHUFFLE DAN GEBRAKAN MENTERI PRABOWO...	@PratamAldi732	Dengan adanya reshuffle ini tentunya rakyat me...	0	2025-09-21 11:38:02
LJIoC4QAP9A	MENAKAR RESHUFFLE DAN GEBRAKAN MENTERI PRABOWO...	@FerliandoTupamahu	Langkah reshuffle udah pas buat perkuat barisan.	0	2025-09-21 11:02:43
LJIoC4QAP9A	MENAKAR RESHUFFLE DAN GEBRAKAN MENTERI PRABOWO...	@CeciliaNaya-d9c8o	Prabowo-Gibran harus terus evaluasi kabinetnya.	0	2025-09-21 09:49:18

3.2. Preprocessing (Pemrosesan awal)

Tahap preprocessing dilakukan untuk memastikan teks komentar siap dianalisis menggunakan metode NLP dan topic modeling. Tahap ini terdiri dari beberapa proses berikut:



Gambar 2. Arsitektur Model LDA

a. Case Folding

Seluruh karakter pada komentar diubah menjadi huruf kecil agar konsisten dan tidak menimbulkan variasi makna akibat perbedaan kapitalisasi [38].

b. Cleaning Data

Pada tahap ini dilakukan pembersihan teks dengan menghapus seluruh elemen yang tidak relevan seperti URL, angka, mention, tag, tanda baca berlebih, emoji, karakter non-ASCII, serta spasi ganda [39].

c. Normalization

Normalisasi dilakukan dengan menyelaraskan kata-kata yang memiliki bentuk tidak baku atau variasi penulisan [40]. Contohnya seperti :
gk → nggak, ngga → tidak, menkeu → menteri keuangan.

Normalisasi membantu mengurangi ketidakteraturan teks yang lazim ditemukan pada komentar YouTube.

d. Tokenization

Komentar yang telah dibersihkan dipecah menjadi unit kata atau token. Proses tokenisasi ini memungkinkan analisis kata per kata dalam model topic modeling [41].

e. Stopword Removal

Kata-kata umum yang tidak memiliki makna signifikan dalam analisis, seperti yang, dan, di, atau, dihapus menggunakan daftar stopwords Bahasa Indonesia yang telah diperluas sesuai karakteristik komentar YouTube [42].

f. Stemming

Setiap kata dikembalikan ke bentuk dasarnya menggunakan stemmer Bahasa Indonesia seperti Sastrawi. Langkah ini bertujuan mengurangi variasi kata sehingga distribusi kata lebih stabil untuk pemodelan [43].

3.3. Build Models (Membangun Model)

Setelah preprocessing selesai, data teks direpresentasikan ke dalam bentuk numerik menggunakan CountVectorizer, representasi ini menghasilkan matriks dokumen kata yang diperlukan oleh model [44]. Algoritma Latent Dirichlet Allocation (LDA) digunakan untuk mengidentifikasi topik-topik tersembunyi dalam komentar. Pada tahap ini, hyperparameter seperti jumlah topik, alpha, beta, dan jumlah iterasi disesuaikan untuk mendapatkan

hasil pemodelan terbaik [45]. LDA dipilih karena kemampuannya memetakan distribusi kata dalam dokumen secara tidak terawasi, sesuai kebutuhan analisis topik.

3.4. Evaluation (Evaluasi)

Evaluasi dilakukan untuk menilai kualitas topik yang dihasilkan LDA. Metrik utama yang digunakan adalah topic diversity, yang mengukur keberagaman kata unik pada setiap topik. Topik dianggap baik jika memiliki kata kunci yang tidak tumpang tindih dengan topik lainnya. Selain itu, digunakan pula coherence score untuk mengevaluasi keterkaitan semantik antar kata dalam setiap topik. Nilai koherensi yang tinggi menunjukkan bahwa topik lebih mudah dipahami oleh manusia [46].

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pemodelan Topik

Pemodelan topik menggunakan algoritma Latent Dirichlet Allocation (LDA) diterapkan pada 51.550 komentar YouTube yang telah melalui tahap preprocessing. Representasi teks dilakukan menggunakan CountVectorizer, sehingga setiap komentar diperlakukan sebagai dokumen yang terdiri dari distribusi kata. Berdasarkan proses evaluasi menggunakan coherence score dan topic diversity, jumlah topik optimal yang dipilih dalam penelitian ini adalah tiga topik.

Tabel 2. LDA Model Result

Topik	Probabilitas
1	0.044*"menteri" + 0.037*"rakyat" + 0.026*"srimulyani" + 0.025*"semoga" + 0.023*"indonesia" + 0.019*"uang" + 0.018*"prabowo" + 0.017*"presiden" + 0.014*"negara" + 0.012*"kerja"
2	0.027*"korupsi" + 0.023*"pajak" + 0.017*"negara" + 0.014*"rakyat" + 0.011*"dpr" + 0.010*"uang" + 0.009*"tambah" + 0.008*"kementerian" + 0.008*"hukum" + 0.007*"hutang"
3	0.098*"ganti" + 0.049*"menteri" + 0.031*"kapolri" + 0.027*"bahlil" + 0.023*"prabowo" + 0.022*"jokowi" + 0.017*"budiarie" + 0.014*"luhut" + 0.013*"reshuffle" + 0.012*"kabinet"

Pemilihan tiga topik didasarkan pada pertimbangan keseimbangan antara keterpahaman topik dan minimnya tumpang tindih kata antar topik. Jumlah topik yang lebih sedikit cenderung menghasilkan tema yang terlalu umum, sedangkan jumlah topik yang lebih banyak menyebabkan fragmentasi topik dan menurunkan interpretabilitas. Dengan demikian, tiga topik dinilai mampu merepresentasikan struktur diskusi publik secara komprehensif.

Setiap topik direpresentasikan oleh sepuluh kata kunci dengan probabilitas tertinggi. Kata-kata tersebut mencerminkan tema dominan yang paling sering

muncul dalam diskusi publik terkait reshuffle Menteri Keuangan 2025.

4.2. Distribusi dan Karakteristik Topik

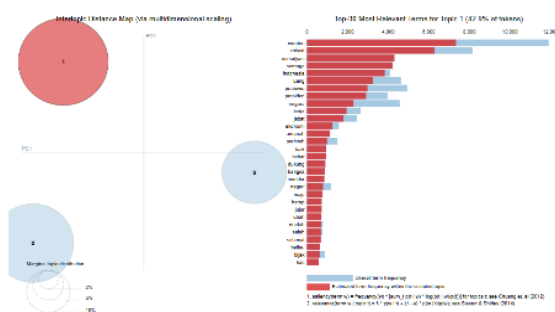
Hasil pemodelan menunjukkan bahwa ketiga topik memiliki karakteristik diskusi yang berbeda, baik dari segi fokus isu maupun orientasi persepsi publik. Distribusi topik menunjukkan bahwa komentar publik tidak terpusat pada satu aspek saja, melainkan terbagi ke dalam dimensi kepemimpinan dan harapan publik, kritik terhadap tata kelola keuangan, serta dinamika politik dan reshuffle kabinet.

Pemisahan topik ini mengindikasikan bahwa wacana reshuffle Menteri Keuangan dipahami publik secara multidimensional. Publik tidak hanya merespons dari sudut pandang ekonomi, tetapi juga mengaitkannya dengan persoalan politik, integritas lembaga negara, serta kepemimpinan nasional secara umum.

4.3. Topik 1: Harapan Publik terhadap Kepemimpinan Menteri Keuangan

Kata kunci dominan:

menteri, rakyat, srimulyani, semoga, indonesia, uang, prabowo, presiden, negara, kerja



Gambar 3. Visualisasi Hasil Pemodelan Topik 1

Topik pertama merepresentasikan persepsi publik yang bersifat normatif dan aspiratif. Dominasi kata “*semoga*” menunjukkan bahwa sebagian besar komentar dalam topik ini mengandung harapan dan doa terhadap kinerja Menteri Keuangan pasca-reshuffle. Kehadiran kata “*rakyat*” dan “*negara*” menandakan bahwa publik menilai kebijakan keuangan negara dari perspektif kesejahteraan masyarakat luas.

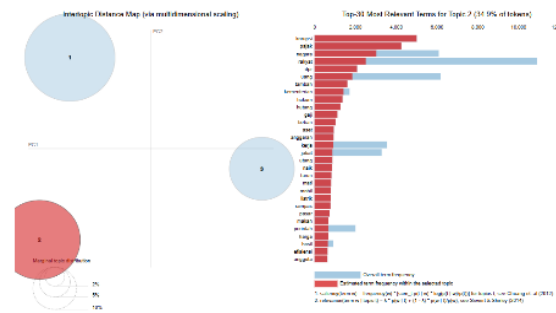
Kemunculan nama “*Sri Mulyani*” mengindikasikan bahwa figur Menteri Keuangan sebelumnya masih menjadi simbol kuat dalam diskursus publik. Hal ini menunjukkan adanya *benchmarking* publik terhadap rekam jejak kepemimpinan yang dianggap berhasil dalam menjaga stabilitas fiskal. Selain itu, keterkaitan antara kata “*presiden*” dan “*prabowo*” memperlihatkan bahwa publik memandang reshuffle Menteri Keuangan sebagai bagian dari kebijakan strategis presiden dalam membangun arah ekonomi nasional.

Secara konseptual, topik ini mencerminkan bentuk kepercayaan bersyarat publik terhadap pemerintah. Dukungan yang diberikan tidak bersifat mutlak, melainkan disertai ekspektasi agar kebijakan yang diambil benar-benar berpihak pada kepentingan rakyat dan keberlanjutan ekonomi Indonesia.

4.4. Topik 2: Kritik terhadap Tata Kelola Keuangan dan Isu Korupsi

Kata kunci dominan:

korupsi, pajak, negara, rakyat, dpr, uang, tambah, kementerian, hukum, hutang



Gambar 4. Visualisasi Hasil Pemodelan Topik 2

Topik kedua menunjukkan sentimen publik yang lebih kritis dan skeptis. Dominasi kata “*korupsi*” menjadi indikator kuat bahwa isu integritas dan transparansi pengelolaan keuangan negara merupakan perhatian utama masyarakat. Publik tidak hanya mengaitkan reshuffle dengan pergantian individu, tetapi juga sebagai peluang untuk memperbaiki sistem yang dianggap bermasalah.

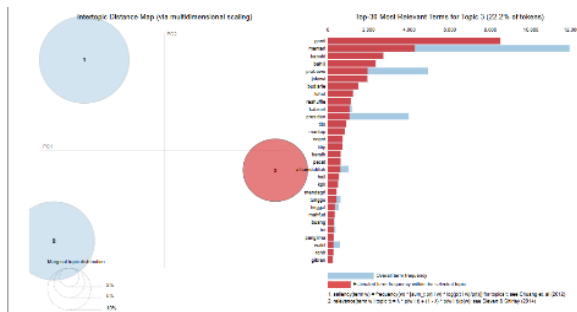
Kata “*pajak*” dan “*hutang*” mencerminkan kekhawatiran publik terhadap kebijakan fiskal yang berdampak langsung pada kehidupan ekonomi masyarakat. Dalam konteks ini, publik menyoroti beban pajak dan utang negara sebagai konsekuensi dari lemahnya tata kelola keuangan. Kehadiran kata “*dpr*” dan “*hukum*” menunjukkan bahwa kritik tidak hanya ditujukan kepada kementerian, tetapi juga pada lembaga legislatif dan sistem penegakan hukum yang dianggap kurang optimal dalam menjalankan fungsi pengawasan.

Topik ini memperlihatkan meningkatnya kesadaran politik dan literasi ekonomi publik di ruang digital. Diskursus yang muncul tidak lagi bersifat emosional semata, melainkan mencerminkan pemahaman struktural terhadap permasalahan keuangan negara.

4.5. Topik 3: Dinamika Politik dan Isu Reshuffle Kabinet

Kata kunci dominan:

ganti, menteri, kapolri, bahlil, prabowo, jokowi, budiarie, luhut, reshuffle, kabinet



Gambar 5. Visualisasi Hasil Pemodelan Topik 3

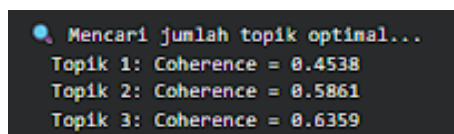
Topik ketiga memiliki nilai coherence tertinggi dan berfokus pada dinamika politik kekuasaan. Kata “ganti” dan “reshuffle” mendominasi diskusi, menandakan bahwa publik sangat memperhatikan proses pergantian jabatan sebagai bagian dari strategi politik pemerintahan.

Munculnya berbagai nama tokoh politik seperti Jokowi, Prabowo, Luhut, dan Bahlil menunjukkan bahwa diskusi publik meluas dari isu Menteri Keuangan ke konfigurasi kabinet secara keseluruhan. Hal ini mengindikasikan bahwa reshuffle dipersepsikan sebagai sinyal perubahan arah kebijakan dan konsolidasi kekuasaan.

Topik ini juga mencerminkan adanya diskursus spekulatif, di mana publik mencoba menafsirkan motif politik di balik reshuffle. Dengan demikian, topik ini menunjukkan bahwa ruang komentar YouTube berfungsi sebagai arena diskusi politik informal yang aktif dan dinamis.

4.6. Evaluasi Coherence Score dan Topic Diversity

Hasil evaluasi menunjukkan bahwa nilai coherence untuk masing-masing topik adalah 0,4538 (Topik 1), 0,5861 (Topik 2), dan 0,6359 (Topik 3). Nilai tertinggi pada Topik 3 menunjukkan bahwa topik ini memiliki keterkaitan semantik yang paling kuat antar kata kunci, sehingga mudah dipahami secara konseptual.



Gambar 6. Hasil Coherence Score

*** Topic Diversity: 0.833

Gambar 7. Hasil Diversity Score

Selain itu, nilai topic diversity sebesar 0,833 menunjukkan bahwa topik-topik yang dihasilkan memiliki tingkat keberagaman kata yang tinggi. Artinya, model berhasil meminimalkan penggunaan kata yang sama pada topik berbeda, sehingga pemisahan tema menjadi lebih jelas dan terstruktur.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa pemodelan topik menggunakan algoritma Latent Dirichlet Allocation (LDA) mampu mengungkap struktur tematik dalam komentar YouTube terkait isu reshuffle Menteri Keuangan tahun 2025 secara efektif. Analisis terhadap 51.550 komentar yang dikumpulkan dari 437 video YouTube menunjukkan bahwa persepsi publik terhadap reshuffle tidak bersifat tunggal, melainkan terbagi ke dalam beberapa dimensi utama yang saling berkaitan.

Hasil pemodelan menghasilkan tiga topik utama yang merepresentasikan pola diskusi dominan publik. Topik pertama menggambarkan harapan dan aspirasi publik terhadap kepemimpinan Menteri Keuangan dan pemerintah secara umum. Pada topik ini, publik mengekspresikan ekspektasi agar kebijakan fiskal yang diambil dapat berpihak pada kesejahteraan rakyat dan menjaga stabilitas ekonomi nasional. Kehadiran figur Menteri Keuangan sebelumnya dalam diskursus publik menunjukkan bahwa rekam jejak kepemimpinan masih menjadi tolok ukur utama dalam penilaian masyarakat.

Topik kedua merefleksikan sikap kritis dan skeptis masyarakat terhadap tata kelola keuangan negara. Isu-isu seperti korupsi, pajak, utang negara, serta lemahnya penegakan hukum menjadi perhatian utama publik. Hal ini mengindikasikan bahwa reshuffle dipandang sebagai momentum penting untuk melakukan evaluasi struktural terhadap sistem pengelolaan keuangan negara, bukan sekadar pergantian aktor. Temuan ini menunjukkan meningkatnya kesadaran publik terhadap transparansi, akuntabilitas, dan integritas lembaga negara.

Topik ketiga berfokus pada dinamika politik dan proses reshuffle kabinet secara keseluruhan. Diskusi publik pada topik ini menyoroti pergantian jabatan sebagai bagian dari strategi politik dan konsolidasi kekuasaan. Munculnya berbagai nama tokoh politik menandakan bahwa publik memahami reshuffle sebagai proses politik yang kompleks dan memiliki implikasi terhadap arah kebijakan pemerintahan ke depan. Topik ini juga memiliki nilai coherence tertinggi, yang menunjukkan bahwa diskursus politik publik cenderung lebih terstruktur dan konsisten secara semantik.

Dari sisi evaluasi model, nilai coherence score yang berada pada rentang moderat hingga tinggi serta nilai topic diversity sebesar 0,833 menunjukkan bahwa topik-topik yang dihasilkan memiliki keterkaitan makna yang baik dan tingkat keberagaman kata yang tinggi. Hal ini menandakan bahwa model LDA yang digunakan berhasil memisahkan tema-tema diskusi secara jelas dan minim tumpang tindih, sehingga layak dijadikan dasar dalam analisis persepsi publik berbasis data teks.

Secara keseluruhan, penelitian ini menegaskan bahwa komentar YouTube dapat dimanfaatkan sebagai sumber data yang relevan dan representatif

untuk memahami persepsi publik terhadap isu politik strategis. Pendekatan topic modeling berbasis LDA terbukti mampu mengungkap pola komunikasi publik yang tidak mudah teridentifikasi melalui analisis manual, sekaligus memberikan gambaran komprehensif mengenai dinamika opini masyarakat di ruang publik digital.

Namun demikian, penelitian ini memiliki beberapa keterbatasan yang perlu diakui. Pertama, dataset hanya mencakup rentang waktu 8-22 September 2025, sehingga menghasilkan snapshot persepsi yang tidak menangkap dinamika jangka panjang sebelum dan sesudah reshuffle. Kedua, penelitian terbatas pada satu platform (YouTube), padahal diskusi publik juga aktif di Twitter/X, TikTok, dan Instagram. Ketiga, preprocessing bahasa Indonesia masih menghadapi tantangan dalam menangani slang intensif dan variasi dialek lokal, meskipun sudah menggunakan Sastrawi stemmer. Keempat, model LDA memiliki keterbatasan dalam menangkap konteks pragmatis dan ironi dalam komunikasi digital. Kelima, coherence score moderat pada Topik 1 (0,4538) mengindikasikan konsistensi semantik yang lebih rendah.

DAFTAR PUSTAKA

- [1] H. Jin, "From Information to Affirmation: An Investigation on the Echo Chamber Effect from YouTube Comments under Technology Product Reviews," arXiv e-prints, pp. 1–13, 2022, doi: <https://doi.org/10.48550/arXiv.2211.00483>.
- [2] D. S. Rahmawati, "Analyzing YouTube Comments on Predatory Pricing: A Digital Citizenship on Netnography Study," vol. 18, no. 1, pp. 12–26, 2025, doi: [10.29313/mediator.v18i1.3347](https://doi.org/10.29313/mediator.v18i1.3347).
- [3] A. Khaidar, "Analisis sentimen di Instagram terhadap Menteri Keuangan Purbaya Yudhi Sadewa menggunakan metode logistic regression," *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 3, pp. 1674–1683, 2025, doi: <https://doi.org/10.23960/jitet.v13i3S1.8002>.
- [4] S. De Putri and A. Toni, "Analisis percakapan netizen pada channel YouTube @TotalPolitik sebagai media ruang publik komunikasi politik," *Sociol. J. Ilm. Kaji. Ilmu Sos. dan Budaya*, vol. 26, no. 2, pp. 133–150, 2024, doi: <https://doi.org/10.23960/sosiologi.v26i2.1389>.
- [5] C. Andrew, A. Budiyantra, and I. Lewenusa, "Analisis sentimen komentar pengguna YouTube untuk ponsel di Indonesia berbasis web," *Sci. J. Ilm. Sains Dan Teknol.*, vol. 3, no. 2, pp. 175–188, 2024.
- [6] N. Royani, C. E. Widodo, and B. Warsito, "Topic Modelling Latent Dirichlet Allocation untuk Klasifikasi Komentar pada Layanan Streaming Platform," *J. Sains dan Teknol.*, vol. 12, no. 3, pp. 815–822, 2023, doi: <https://doi.org/10.23887/jstundiksha.v12i3.68492>.
- [7] D. Prastyo, I. H. Mursyidin, and D. Irawan, "Klasifikasi Sentimen Komentar YouTube dengan NLP pada Debat Pilkada Banten 2024," *bit-Tech*, vol. 7, no. 2, pp. 413–421, 2024, doi: [10.32877/bt.v7i2.1833](https://doi.org/10.32877/bt.v7i2.1833).
- [8] H. Jin, "From Information to Affirmation: An Investigation on the Echo Chamber Effect from YouTube Comments under Technology Product Reviews," arXiv e-prints, pp. 1–13, 2022, doi: <https://doi.org/10.48550/arXiv.2211.00483>.
- [9] Statista, "YouTube—Statistics & Facts," 2024. [Online]. Available: <https://www.statista.com/topics/2019/youtube/>. [Accessed: Dec. 2024].
- [10] D. S. Rahmawati, "Analyzing YouTube Comments on Predatory Pricing: A Digital Citizenship on Netnography Study," *Mediator J. Commun. Stud.*, vol. 18, no. 1, pp. 12–26, 2025, doi: [10.29313/mediator.v18i1.3347](https://doi.org/10.29313/mediator.v18i1.3347).
- [11] Z. Papacharissi, *Affective Publics: Sentiment, Technology, and Politics*. Oxford: Oxford University Press, 2015.
- [12] D. A. Yates, C. Wagner, and A. Majchrzak, "Hope and Aspiration in Democratic Discourse Online," *New Media Soc.*, vol. 26, no. 4, pp. 1856–1872, 2024.
- [13] B. D. Loader and D. Mercea, "Networking Democracy? Information, Communication & Society," *Inf. Commun. Soc.*, vol. 14, no. 6, pp. 757–769, 2011.
- [14] D. Prastyo, I. H. Mursyidin, and D. Irawan, "Klasifikasi Sentimen Komentar YouTube dengan NLP pada Debat Pilkada Banten 2024," *bit-Tech*, vol. 7, no. 2, pp. 413–421, 2024, doi: [10.32877/bt.v7i2.1833](https://doi.org/10.32877/bt.v7i2.1833).
- [15] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol: O'Reilly Media, 2009.
- [16] T. S. Hapsari and Y. Nataliani, "Text Mining with Latent Dirichlet Allocation for Analyzing Public Comments on the M-Passport Application," *J. Tek. Inform.*, vol. 5, no. 4, pp. 1093–1101, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.4.1929>.
- [17] R. Hidayat et al., "Indonesian Natural Language Processing: A Survey," arXiv preprint arXiv:2311.09173, 2023. [Online]. Available: <https://arxiv.org/abs/2311.09173>.
- [18] Sastrawi, "Indonesian Morphological Analyzer (Sastrawi)," GitHub, 2024. [Online]. Available: <https://github.com/sastrawi/sastrawi>. [Accessed: Dec. 2024].
- [19] F. Koto et al., "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," arXiv preprint arXiv:2009.05141, 2020. [Online]. Available: <https://arxiv.org/abs/2009.05141>.

- [20] N. Royani, C. E. Widodo, and B. Warsito, "Topic Modelling Latent Dirichlet Allocation untuk Klasifikasi Komentar pada Layanan Streaming Platform," *J. Sains dan Teknol.*, vol. 12, no. 3, pp. 815–822, 2023, doi: <https://doi.org/10.23887/jstundiksha.v12i3.68492>.
- [21] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
- [22] M. Hoffman, F. R. Bach, and D. M. Blei, "Online Learning for Latent Dirichlet Allocation," in *Advances in Neural Information Processing Systems* 23, 2010, pp. 856–864.
- [23] M. Grootendorst, "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure," *arXiv preprint arXiv:2203.05556*, 2022. [Online]. Available: <https://arxiv.org/abs/2203.05556>.
- [24] D. Angelov, "Top2Vec: Distributed Representations of Topics," *arXiv preprint arXiv:2008.09470*, 2020. [Online]. Available: <https://arxiv.org/abs/2008.09470>.
- [25] A. Khaidar, "Analisis Sentimen di Instagram terhadap Menteri Keuangan Purbaya Yudhi Sadewa Menggunakan Metode Logistic Regression," *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 3, pp. 1674–1683, 2025, doi: <https://doi.org/10.23960/jitet.v13i3S1.8002>.
- [26] M. Röder, A. Both, and A. Hinneburg, "Exploring the Space of Topic Coherence Measures," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 2015, pp. 399–408.
- [27] H. M. Wallach et al., "Evaluation Methods for Topic Models," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 1105–1112.
- [28] J. Chang, J. Boyd-Graber, S. Gerrish, C. Wang, and D. M. Blei, "Reading Tea Leaves: How Humans Interpret Topic Models," in *Advances in Neural Information Processing Systems* 22, 2009, pp. 288–296.
- [29] F. Sugandi, "Analisis Sentiment Masyarakat Indonesia pada Media Sosial terhadap Isu Ijazah Palsu Mantan Presiden Menggunakan Algoritma Berbasis Transformer (BERT)," *J. Soc. Sci. Res.*, vol. 8, no. 3, pp. 4762–4768, 2025, doi: <https://doi.org/10.54314/jssr.v8i3.4209>.
- [30] V. Alpiana, A. Salam, F. Alzami, I. Rizqa, and D. Aqmal, "Analisis Topic-Modelling Menggunakan Latent Dirichlet Allocation (LDA) Pada Ulasan Sosial Media YouTube," *J. Media Inform. BUDIDARMA*, vol. 8, no. 1, pp. 332–341, 2024, doi: <https://doi.org/10.30865/mib.v8i1.7127>.
- [31] *Topic Modelling Latent Dirichlet Allocation untuk Klasifikasi Komentar pada Layanan Streaming Platform*, *Jurnal Sains dan Teknologi*, vol. 12, no. 3, pp. 815–822, 2023, doi: [10.23887/jstundiksha.v12i3.68492](https://doi.org/10.23887/jstundiksha.v12i3.68492).
- [32] "Penerapan Pemodelan Topik Komentar Melalui Media Sosial Twitter Menggunakan Latent Dirichlet Allocation (Studi Kasus: Pemerintah Kota Malang)," *J-PTIUK*, vol. 8, no. 7, Jul. 2024. [Online]. Available: <https://j-ptiuk.ub.ac.id/index.php/j-ptiuk/article/view/13979>. Accessed: Dec. 22, 2025.
- [33] E. S. . Negara and D. . Triadi, "Topic modeling using latent dirichlet allocation (LDA) on twitter data with Indonesia keyword", *Bull.Socinf.The.App*, vol. 5, no. 2, pp. 124–132, Jan. 2022.
- [34] "ANALISIS SENTIMEN DAN PEMODELAN TOPIK PADA TWEET TERKAIT DATA BADAN PUSAT STATISTIK", *j.sist.inf.inform.*, vol. 6, no. 2, pp. 165–176, Aug. 2023, doi: [10.47080/simika.v6i2.2789](https://doi.org/10.47080/simika.v6i2.2789).
- [35] M. R. Al Basyary, "YouTube Comment Sentiment Analysis on Deddy Corbuzier and BEM UI's Podcast Using TF-IDF and Naïve Bayes," *SAINTEKBU J. Sci. Technol.*, vol. 16, no. 02, pp. 10–17, 2024.
- [36] E. D. Absharina and F. Nurqolbiah, "Analisis Sentimen Komentar Video Youtube Review Starlink Indonesia oleh Gadgetin menggunakan VADER dan TextBlob," *Indones. J. Comput. Sci.*, vol. 14, no. 3, pp. 4383–4389, 2025, doi: <https://doi.org/10.33022/ijcs.v14i3.4686>.
- [37] P. D. N. Sari, Y. T. Widayati, and S. A. Prakoso, "Analysis of Public Sentiment Towards the Use of AI in Monitoring Waste via the SEMAR Monitoring Web and Its Impact on Flood Management in Semarang City," *J. Teknologi Inform. dan Komput. MH. Thamrin*, vol. 11, no. 2, pp. 805–824, 2025, doi: <https://doi.org/10.37012/jtik.v11i2.2819>.
- [38] T. S. Hapsari and Y. Nataliani, "Text mining with Latent Dirichlet Allocation for analyzing public comments on the M-Passport application," **J. Tek. Inform.**, vol. 5, no. 4, pp. 1093–1101, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.4.1929>.
- [39] L. Septiana, V. Yasin, and A. Z. Sianipar, "Analyzing public sentiment on youtube comments regarding the free lunch policy using the Support Vector Machine (SVM) algorithm," *J. Mandiri IT*, vol. 14, no. 1, pp. 129–138, 2025, doi: <https://doi.org/10.35335/mandiri.v14i1.434>.
- [40] F. Sugandi, "Analisis sentiment masyarakat Indonesia pada media sosial terhadap isu ijazah palsu mantan Presiden menggunakan algoritma berbasis Transformer (BERT)," **J. Soc. Sci. Res.**, vol. 8, no. 3, pp. 4762–4768, 2025, doi: <https://doi.org/10.54314/jssr.v8i3.4209>.
- [41] G. Airlangga, "Spam Detection in YouTube Comments Using Deep Learning Models: A

- Comparative Study of MLP , CNN , LSTM , BiLSTM , GRU , and Attention Mechanisms,” MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 4, no. 4, pp. 1533–1538, 2024, doi: <https://doi.org/10.57152/malcom.v4i4.167>.
- [42] 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, pp. 326–334, 2023.
- [43] V. Alpiana, A. Salam, F. Alzami, I. Rizqa, and D. Aqmala, “Analisis Topic-Modelling Menggunakan Latent Dirichlet Allocation (LDA) Pada Ulasan Sosial Media Youtube,” J. MEDIA Inform. BUDIDARMA, vol. 8, no. 1, pp. 332–341, 2024, doi: <https://doi.org/10.30865/mib.v8i1.7127>.
- [44] F. Panjaitan, “Perbandingan penggunaan TfIdfVectorizer, CountVectorizer, dan HashingVectorizer dengan optimalisasi parameter pada machine learning untuk analisis sentimen Pemilu 2024,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 7413–7419, 2024, doi: <https://doi.org/10.36040/jati.v8i4.10288>.
- [45] L. J. Hagg et al., “Examining Analytic Practices in Latent Dirichlet Allocation Within Psychological Science: Scoping Review,” J. Med. Internet Res., vol. 24, no. 11, pp. 1–24, 2022, doi: [10.2196/33166](https://doi.org/10.2196/33166).
- [46] W. Wahyuni, T. P. Lestari, M. Apriliana, and R. Gumelta, “Implementation of BERTopic for Topic Modeling Analysis of the Free Nutritious Meal Program Based on YouTube Comments,” J. Appl. Informatics Comput., vol. 9, no. 4, pp. 1964–1971, 2025, doi: <https://doi.org/10.30871/jaic.v9i4.9754>.