Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Analisis Perbandingan Performa NMF dengan LDA pada Topik Modeling Berita Online Indonesia

Latifah Nurrohmah Handayani*, Lucia Nugraheni Harnaningrum
Teknologi Informasi, Magister Teknologi Informasi, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia
Email: [1,*]latifah.hand99@gmail.com, [2]ningrum@utdi.ac.id
Submitted: **12/01/2026**; Accepted: **13/02/2026**; Published: **31/03/2026**

Abstrak– Pertumbuhan konten berita digital di Indonesia menciptakan tantangan dalam mengekstraksi topik-topik utama dari dataset berskala besar secara otomatis. Permasalahan utama adalah belum adanya kajian empiris komparatif yang secara khusus membandingkan metode *topic modeling* NMF dan LDA pada korpus berita *online* Indonesia. Penelitian ini bertujuan melakukan analisis komparatif performa *Non-negative Matrix Factorization* (NMF) dan *Latent Dirichlet Allocation* (LDA) dalam pemodelan topik berita *online* Indonesia, serta memetakan pola pemberitaan dari tiga media nasional: CNBC Indonesia, Kompas.com, dan Detik.com. Dataset terdiri dari 4.500 artikel berita dengan *preprocessing* meliputi tokenisasi, penghapusan *stopwords*, serta ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk NMF dan *Count Vectorizer* untuk LDA. Evaluasi performa menggunakan *coherence score* (C_v), *topic diversity*, *silhouette score*, dan uji *chi-square*. Hasil analisis menunjukkan NMF memberikan kualitas topik dan efisiensi komputasi yang lebih baik, ditunjukkan oleh nilai *coherence score* 34,7% lebih tinggi, *topic diversity* 11.9% lebih baik, *silhouette score* 48% lebih besar, serta waktu *training* (1.36 *seconds* vs. 86.11 *seconds*) yang 63,3× lebih cepat dibandingkan LDA. Analisis mengidentifikasi 10 topik utama dan mengonfirmasi perbedaan fokus editorial yang signifikan antar media ($p < 0.001$), dengan CNBC Indonesia mendominasi topik ekonomi-keuangan, sementara Kompas.com dan Detik.com lebih fokus pada isu sosial-kemasyarakatan. Kontribusi utama penelitian ini adalah menyediakan bukti empiris mengenai performa komparatif NMF dan LDA pada korpus berita Indonesia serta memetakan diferensiasi konten media nasional. Hasil penelitian menunjukkan bahwa NMF lebih sesuai digunakan untuk pemodelan topik berita Indonesia yang menekankan interpretabilitas topik dan efisiensi komputasi.

Kata Kunci: Topic Modeling; Non-negative Matrix Factorization; Latent Dirichlet Allocation; Natural Language Processing; Analisis Berita

Abstract– The growth of digital news content in Indonesia presents challenges in automatically extracting main topics from large-scale datasets. A key issue is the lack of empirical studies specifically comparing Non-negative Matrix Factorization (NMF) and Latent Dirichlet Allocation (LDA) on Indonesian online news corpora. This study aims to conduct a comparative analysis of NMF and LDA performance for topic modeling of Indonesian online news and to map reporting patterns across three national media outlets: CNBC Indonesia, Kompas.com, and Detik.com. The dataset consists of 4,500 news articles, with preprocessing including tokenization, stopword removal, and feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF) for NMF and Count Vectorizer for LDA. Performance was evaluated using coherence score (C_v), topic diversity, silhouette score, and chi-square test. Results indicate that NMF delivers better topic quality and computational efficiency, with a 34.7% higher coherence score, 11.9% higher topic diversity, 48% higher silhouette score, and 63.3× faster training time (1.36 seconds vs. 86.11 seconds) compared to LDA. Analysis identified 10 main topics and confirmed significant differences in editorial focus across media ($p < 0.001$), with CNBC Indonesia dominating economic–financial topics, while Kompas.com and Detik.com focused more on social–community issues. The study's main contribution is providing empirical evidence of the comparative performance of NMF and LDA on Indonesian news corpora and mapping content differentiation among national media. Findings suggest that NMF is more suitable for Indonesian news topic modeling, emphasizing topic interpretability and computational efficiency.

Keywords: Topic Modeling; Non-negative Matrix Factorization; Latent Dirichlet Allocation; Natural Language Processing; News Analysis

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong pertumbuhan signifikan dalam produksi konten digital, khususnya berita *online* di Indonesia [1]. Setiap hari, jutaan artikel berita dipublikasikan oleh berbagai portal media, sehingga menimbulkan tantangan dalam mengorganisasi dan mengekstraksi informasi bermakna dari volume data yang besar [2]. Sebagai ilustrasi, tiga media nasional seperti CNBC Indonesia, Kompas.com, dan Detik.com secara kolektif mempublikasikan ratusan artikel berita setiap harinya dengan beragam topik mulai dari ekonomi, politik, sosial, hingga kesehatan. Dataset berita *online* tersebut merepresentasikan ribuan artikel dengan variasi topik lintas media dan periode waktu tertentu. Volume publikasi yang tinggi ini menciptakan permasalahan utama dalam mengidentifikasi tema-tema dominan dan pola pemberitaan secara otomatis tanpa memerlukan analisis manual yang memakan waktu dan tenaga. Namun, hingga saat ini belum terdapat acuan empiris yang jelas mengenai metode *topic modeling* yang paling sesuai untuk karakteristik berita *online* Indonesia, sehingga pemilihan metode masih belum didasarkan pada evaluasi empiris yang konsisten. Dalam konteks ini, *topic modeling* muncul sebagai salah satu teknik *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi struktur tematik dalam kumpulan dokumen berskala besar [3]. Teknik ini memungkinkan ekstraksi tema-tema utama dari kumpulan dokumen teks tanpa memerlukan anotasi manual, sehingga efisien untuk analisis data tekstual [4].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Dua algoritma yang banyak digunakan dalam *topic modeling* adalah *Latent Dirichlet Allocation* (LDA) dan *Non-negative Matrix Factorization* (NMF) [5]. LDA merupakan model *topic modeling* yang mengasumsikan bahwa setiap dokumen dibangun dari perpaduan beberapa topik laten, di mana setiap dokumen memiliki distribusi topik tertentu[6], sedangkan NMF merupakan metode aljabar linear yang mendekomposisi matriks dokumen-term menjadi dua matriks non-negatif sehingga menghasilkan representasi topik yang relatif lebih mudah diinterpretasi [7]. Meskipun kedua metode telah diterapkan dalam berbagai domain, studi komparatif yang secara sistematis membandingkan performa NMF dan LDA pada teks bahasa Indonesia, khususnya berita *online*, masih terbatas [8]. Karakteristik linguistik bahasa Indonesia juga berpotensi memengaruhi kinerja algoritma *topic modeling* [9], sehingga analisis komparatif diperlukan untuk menentukan metode yang lebih sesuai.

Sejumlah penelitian sebelumnya telah membahas *topic modeling* pada berita Indonesia. Egger dan Yu melakukan studi komparatif NMF dan LDA pada teks pendek dan menemukan bahwa NMF menghasilkan topik yang lebih mudah diinterpretasi[1]. Puspita menggunakan LDA untuk pemodelan topik berita *online* dengan evaluasi *coherence score* dan penilaian manusia [3]. Afidh dan Syahrial menerapkan NMF dengan pendekatan *n-gram* pada 53.920 artikel berita Indonesia[8]. Khairani menunjukkan bahwa tahapan *preprocessing* memiliki pengaruh signifikan terhadap performa model pada teks bahasa Indonesia[10]. Studi komparatif terkini oleh Nanyonga dan Babalola melaporkan bahwa NMF cenderung menghasilkan topik yang lebih terpisah dan mudah diinterpretasi pada dataset teks berskala besar, khususnya data berita [11][12].

Meskipun berbagai studi telah dilakukan, beberapa kesenjangan penelitian (*research gap*) masih dapat diidentifikasi. Pertama, analisis komparatif antara NMF dan LDA yang secara khusus diterapkan pada dataset berita *online* Indonesia dengan menggunakan metrik evaluasi yang beragam masih terbatas. Sebagian penelitian sebelumnya umumnya hanya menggunakan satu atau dua metrik evaluasi, sehingga penilaian performa model menjadi kurang menyeluruh. Kedua, kajian yang menganalisis pola distribusi topik antar media berita Indonesia untuk melihat perbedaan fokus pemberitaan masih relatif minim, karena penelitian sebelumnya lebih menekankan aspek teknis algoritma. Ketiga, karakteristik bahasa Indonesia, seperti penggunaan imbuhan dan kata majemuk, dapat memengaruhi kinerja metode *topic modeling*, namun aspek ini masih jarang dibahas dalam studi perbandingan NMF dan LDA pada dataset berskala menengah.

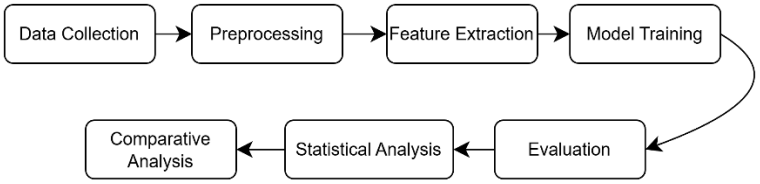
Berdasarkan permasalahan tersebut, penelitian ini membandingkan kinerja NMF dan LDA dalam memodelkan topik berita *online* Indonesia menggunakan 4.500 artikel dari tiga media nasional, yaitu CNBC Indonesia, Kompas.com, dan Detik.com. Analisis dilakukan untuk mengevaluasi performa kedua metode dari berbagai sudut pandang, sekaligus mengidentifikasi topik-topik utama yang muncul serta pola distribusinya di masing-masing media. Dengan demikian ini, penelitian ini tidak hanya menilai aspek teknis model, tetapi juga mengaitkannya dengan karakteristik pemberitaan media.

Hasil penelitian ini diharapkan memberikan gambaran yang lebih jelas mengenai kelebihan dan keterbatasan NMF dan LDA dalam konteks berita Indonesia, serta menjadi acuan dalam pemilihan metode *topic modeling* yang sesuai dengan kebutuhan analisis topik dan interpretasi. Secara praktis, temuan penelitian dapat dimanfaatkan oleh redaksi media untuk memahami pola pemberitaan dan posisi editorial, oleh analis konten dan pemasar untuk menyusun strategi berbasis topik dominan, serta oleh peneliti NLP dan komunikasi sebagai referensi empiris dalam analisis teks berbahasa Indonesia.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan eksperimental komparatif dengan tahapan sistematis yang digambarkan pada Gambar 1. Tahapan dimulai dari pengumpulan data, *preprocessing*, ekstraksi fitur, *training model*, evaluasi, hingga analisis komparatif.



Gambar 1. Diagram Alur Tahapan Penelitian

2.1.1 *Data Collection*

Tahap pertama pada penelitian ini adalah pengumpulan data dari tiga portal berita *online* nasional, yaitu CNBC Indonesia, Kompas.com, dan Detik.com. Pemilihan ketiga media tersebut didasarkan pada perbedaan karakteristik dan segmentasi media, yaitu media bisnis dan ekonomi, media umum berkualitas, serta portal berita cepat dengan volume publikasi tinggi. Selain itu, ketiga portal memiliki tingkat kunjungan (*traffic*) yang besar sehingga dianggap representatif terhadap lanskap pemberitaan daring di Indonesia. Data dikumpulkan menggunakan teknik

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

web scraping dengan total 4.500 artikel berita, masing-masing 1.500 artikel per media, yang dipublikasikan dalam rentang waktu September hingga November 2025, dengan proses pengambilan dilakukan secara bulanan untuk menjaga konsistensi distribusi temporal. Pemilihan artikel dilakukan melalui query filter terstruktur untuk memastikan relevansi terhadap isu kebijakan publik dan pemerintahan. Artikel dipilih berdasarkan kemunculan kata kunci yang berkaitan dengan pemerintahan, kebijakan dan regulasi, keuangan publik, proses legislatif dan pemilu, hukum, serta keamanan nasional, dan diperkuat dengan kata kunci yang merepresentasikan konteks kebijakan. Artikel dengan konten hiburan, olahraga, gaya hidup, konten viral, dan kriminal non-kebijakan dikecualikan dari dataset untuk menjaga fokus analisis. Proses pengumpulan data dilakukan menggunakan platform Apache Solr dengan filter domain dan rentang waktu per bulan, kemudian diekspor dalam format CSV untuk tahap pemrosesan dan analisis selanjutnya.

2.1.2 *Preprocessing*

Pada tahap pra-pemrosesan (preprocessing) ini bertujuan untuk meningkatkan kualitas data sebelum dilakukan pemodelan topik. Tahapan ini meliputi: (1) C*ase folding* untuk mengubah seluruh teks menjadi huruf kecil, (2) tokenisasi untuk memecah teks menjadi unit kata, (3) *Stopword removal* menggunakan daftar *stopwords* bahasa Indonesia yang terdiri dari 758 kata, serta (4) *Filtering* kata dengan panjang minimal empat karakter untuk menghilangkan kata-kata yang kurang informatif. Proses pra-pemrosesan ini bertujuan untuk mengurangi *noise* dan meningkatkan relevansi fitur teks [1], [2]. Teknik-teknik pra-pemrosesan tersebut merupakan bagian dari pendekatan dasar dalam *Natural Language Processing* untuk membentuk korpus teks dan representasi token yang lebih informatif sebelum analisis lebih lanjut, sehingga hasil *topic modeling* menjadi lebih stabil dan mudah diinterpretasikan [13], [14]. Pendekatan ini sejalan dengan praktik umum dalam analisis teks berbahasa Indonesia yang menekankan kesederhanaan dan efisiensi pemrosesan.

2.1.3 *Feature Extraction*

Merupakan tahapan yang dilakukan menggunakan dua pendekatan berbeda sesuai dengan karakteristik masing-masing algoritma. Untuk metode *Non-negative Matrix Factorization* (NMF) digunakan *TF-IDF Vectorizer* dengan parameter *max_df* = 0.95, *min_df* = 3, *max_features* = 5000. TF-IDF (*Term Frequency–Inverse Document Frequency*) merupakan metode pembobotan kata yang mengukur tingkat kepentingan suatu term dalam dokumen berdasarkan frekuensi kemunculannya dan tingkat kelangkaannya dalam kumpulan dokumen, sehingga kata yang lebih spesifik dan membedakan antar dokumen memperoleh bobot yang lebih tinggi [15]. Pendekatan ini konsisten dengan studi yang menerapkan NMF dengan TF-IDF untuk menemukan struktur topik pada korpora teks besar seperti literatur Covid-19 dan menunjukkan bahwa TF-IDF adalah representasi fitur efektif untuk NMF dalam konteks dekomposisi topik[16]. Sementara itu, untuk metode *Latent Dirichlet Allocation* (LDA) digunakan *Count Vectorizer* dengan parameter yang sama. Perbedaan representasi fitur ini didasarkan pada karakteristik algoritma, di mana NMF bekerja lebih optimal dengan representasi TF-IDF, sedangkan LDA sebagai model probabilistik generatif memerlukan data berupa frekuensi kemunculan kata (*raw counts*) sebagai dasar pembentukan distribusi probabilistik dokumen–topik dan topik–kata. Penggunaan *raw count* sejalan dengan asumsi dasar LDA berbasis distribusi multinomial dan *Dirichlet*, serta telah diterapkan secara konsisten dalam berbagai penelitian pemodelan topik pada data teks berbahasa Indonesia [3], [4], [17].

2.1.4 *Model Training*

Penentuan jumlah topik dilakukan menggunakan pendekatan *trial-and-error* terbatas dengan mengevaluasi beberapa konfigurasi jumlah topik pada rentang 5–15 topik. Pendekatan ini dipilih untuk menghindari *over-clustering* yang berpotensi menghasilkan topik kurang informatif dan menurunkan interpretabilitas model, sebagaimana dilaporkan pada penelitian sebelumnya. Berdasarkan hasil eksplorasi awal tersebut, jumlah 10 topik dipilih karena memberikan keseimbangan terbaik antara tingkat detail topik dan keterbacaan hasil analisis [18], [19]. Pelatihan model (*model training*) dilakukan dengan jumlah topik yang ditetapkan sebanyak 10 topik untuk kedua algoritma guna menjaga konsistensi perbandingan. Model NMF dilatih menggunakan parameter *n_components* = 10, *random_state* = 42, *max_iter* = 500, serta metode inisialisasi *NNDSVDa*. Sementara itu, model LDA dilatih dengan parameter *n_components* = 10, *random_state* = 42, *learning_method* = ‘online’, *max_iter* = 50, *batch_size* = 128, dan *learning_decay* = 0.7, parameter yang tidak dispesifikasikan mengikuti konfigurasi default sebagaimana diimplementasikan dalam pustaka *scikit-learn* [20]. Secara operasional, penerapan algoritma dilakukan dengan memanfaatkan representasi fitur hasil tahap ekstraksi untuk membentuk struktur topik laten. Pada metode *Non-negative Matrix Factorization* (NMF), matriks TF-IDF didekomposisi menjadi matriks dokumen–topik dan topik–kata melalui proses optimasi iteratif hingga mencapai kondisi konvergen. Sementara itu, metode *Latent Dirichlet Allocation* (LDA) memodelkan distribusi probabilitas topik pada dokumen dan distribusi kata pada topik menggunakan pendekatan generatif berbasis estimasi iteratif. Seluruh proses pemodelan diterapkan dengan jumlah topik yang sama guna menjaga konsistensi dan keterbandingan hasil analisis.

2.1.5 *Evaluation*

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Evaluasi model ini menggunakan pendekatan multi-metrik untuk penilaian yang menyeluruh. Metrik evaluasi yang digunakan meliputi: (1) *Coherence Score* (C_v) untuk mengukur koherensi semantik topik. Coherence score digunakan sebagai metrik utama karena terbukti lebih konsisten dengan interpretasi manusia dibandingkan perplexity dalam evaluasi kualitas topik[21], (2) *Topic Diversity* untuk menilai tingkat keunikan antar topik, (3) *Silhouette Score* untuk mengukur kualitas pengelompokan dokumen, (4) *Average Topic Probability* untuk menilai tingkat keyakinan penugasan topik, (5) *Training Time* untuk mengevaluasi efisiensi komputasi.

2.1.6 *Statistical Analysis*

Analisis statistik dilakukan terhadap hasil pemodelan topik menggunakan uji *chi-square* untuk menguji independensi distribusi topik terhadap media, serta analisis korelasi *Pearson* untuk mengidentifikasi tren temporal topik selama periode observasi.

2.1.7 *Comparative Analysis*

Analisis komparatif dilakukan untuk membandingkan performa metode *Non-negative Matrix Factorization* (NMF) dan *Latent Dirichlet Allocation* (LDA) dalam pemodelan topik berita *online* Indonesia berdasarkan seluruh metrik evaluasi yang digunakan.

2.2 Dataset dan Sumber Data

Dataset penelitian terdiri dari 4.500 artikel berita *online* Indonesia yang dikumpulkan dari tiga portal media terkemuka dengan distribusi *balanced* (1.500 artikel per media). Karakteristik dataset ditunjukkan pada Tabel 1.

Tabel 1. Karakteristik Dataset

Media	Jumlah Artikel	Periode
CNBC Indonesia	1500	Sep–Nov 2025
Kompas.com	1500	Sep–Nov 2025
Detik.com	1500	Sep–Nov 2025

Dataset mencakup berbagai kategori berita termasuk ekonomi, politik, sosial, hukum, pendidikan, kesehatan, dan internasional. Distribusi temporal yang merata memastikan representasi yang adil dari dinamika pemberitaan selama periode observasi.

2.3 Metode Evaluasi

Evaluasi performa model dilakukan menggunakan pendekatan multi-metrik untuk memberikan penilaian yang menyeluruh terhadap kedua algoritma *topic modeling*.

a) *Coherence Score*

Coherence Score digunakan untuk mengukur kualitas topik berdasarkan keterkaitan semantik antar kata dalam satu topik. Metrik ini merupakan salah satu indikator evaluasi yang umum dalam penelitian *topic modeling* karena memiliki korelasi kuat dengan interpretabilitas topik oleh manusia [22][23]. *Coherence score* mencoba merepresentasikan persepsi kualitas manusia terhadap topik dalam bentuk nilai objektif yang mudah dievaluasi [24].

Coherence Score dihitung menggunakan *Gensim CoherenceModel* dengan persamaan berikut:

$$C_v = \frac{1}{T}\sum_{\{i=1\}}^{\{T\}} coherence(t_i) \quad (1)$$

T adalah jumlah topik, dan *coherence*(t_i) merupakan skor koherensi topik ke-i yang dihitung berdasarkan ko-okurensi kata dalam *sliding window*. Pendekatan berbasis ko-okurensi ini terbukti efektif untuk menilai apakah kata-kata utama dalam suatu topik muncul bersama secara konsisten dalam dokumen yang sama[11]. Sebagai ilustrasi, pada salah satu topik dengan sepuluh kata utama, probabilitas kemunculan bersama pasangan kata *energi* dan *listrik* dari 1.000 dokumen adalah:

$$P(\text{energi,listrik}) = \frac{45}{1000} = 0.045$$

Dengan probabilitas individual:

$$P(\text{energi}) = 0.120, P(\text{listrik}) = 0.095$$

Nilai *Normalized Pointwise Mutual Information* (NPMI) dihitung sebagai:

$$\text{NPMI} = \frac{\log\left(\frac{P(w_1,w_2)}{P(w_1)P(w_2)}\right)}{-\log P(w_1,w_2)} = 0.712$$

Untuk sepuluh kata terdapat $\binom{10}{2} = 45$ pasangan kata, sehingga diperoleh $C_v(\text{Topik}) = 0.754$. Nilai *Coherence Score* model diperoleh dari rata-rata seluruh topik.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

b) *Topic Diversity*
Topic Diversity merupakan metrik yang mengukur seberapa beragam topik-topik yang dihasilkan oleh model *topic modeling* [25]. Metrik ini dihitung sebagai proporsi kata unik pada top-N kata dari seluruh topik :

$$Diversity = \frac{|unique_{words}|}{|total_{words}|} \quad (2)$$

unique_words adalah himpunan kata unik dari seluruh topik, sedangkan *total_words* adalah total kata dari top-N kata pada semua topik. Nilai *diversity* yang tinggi menunjukkan bahwa model mampu menghasilkan topik-topik yang tidak redundan dan memiliki karakteristik yang berbeda satu sama lain[26]. Sebagai ilustrasi, dari 100 kata teratas (10 topik × 10 kata), model NMF menghasilkan 94 kata unik dan LDA menghasilkan 84 kata unik, sehingga diperoleh:

$$\text{Diversity(NMF)} = 0.9400, \text{Diversity(LDA)} = 0.8400$$

c) *Silhouette Score*
Silhouette Score merupakan metrik *unsupervised* yang digunakan untuk mengukur kualitas hasil *clustering* tanpa memerlukan data *training*[27]. *Silhouette Score* didefinisikan sebagai:

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \quad (3)$$

a(i) adalah rata-rata jarak data ke anggota lain dalam klaster yang sama (intra-cluster), dan b(i) adalah rata-rata jarak data ke klaster terdekat lainnya (*inter-cluster*). Nilai *silhouette score* berkisar antara -1 hingga +1, di mana nilai di atas 0.5 dianggap "*reasonable*" dan nilai di atas 0.7 dianggap "*strong*" untuk kualitas *clustering* [28]. Sebagai ilustrasi, untuk sebuah dokumen dengan $a(i) = 0.856$dan $b(i) = 0.892$, diperoleh:

$$s(i) = \frac{0.892 - 0.856}{0.892} = 0.040$$

Nilai *Silhouette Score* model dihitung sebagai rata-rata nilai $s(i)$seluruh dokumen.

d) Uji *Chi-Square* (χ^2)
Uji *chi-square* merupakan uji statistik yang digunakan untuk menguji independensi antara dua variabel kategorikal[11]. Dalam konteks penelitian ini, uji *chi-square* digunakan untuk menguji independensi distribusi topik terhadap media dengan persamaan:

$$\chi^2 = \sum \frac{(O_{\{ij\}} - E_{\{ij\}})^2}{E_{\{ij\}}} \quad (4)$$

O_{ij} adalah frekuensi observasi dan E_{ij} adalah frekuensi ekspektasi dari distribusi topik terhadap media. Nilai *p-value* yang dihasilkan dari uji ini menunjukkan signifikansi statistik dari perbedaan distribusi topik antar media.

3. HASIL DAN PEMBAHASAN

3.1 Perbandingan Metrik Evaluasi

Hasil evaluasi komprehensif menunjukkan perbedaan signifikan antara performa NMF dan LDA pada berbagai metrik. Tabel 2 menyajikan ringkasan perbandingan metrik evaluasi.

Tabel 2. Perbandingan Komprehensif Metrik Evaluasi

Metrik	NMF	LDA	Unggul	Selisih
Coherence Score (C_v)	0.7544	0.5600	NMF	+34.7%
Topic Diversity	0.9400	0.8400	NMF	+11.9%
Silhouette Score	0.0339	0.0229	NMF	+48.0%
Avg Topic Probability	0.0851	0.5587	LDA	-84.8%
Training Time (seconds)	1.36	86.11	NMF	63.3× lebih cepat

Berdasarkan Tabel 2, NMF Nilai *coherence* NMF yang 34.7% lebih tinggi, yaitu 0.7544 dibandingkan 0.5600 untuk LDA, menunjukkan bahwa topik yang dihasilkan NMF jauh lebih mudah dipahami. Dalam praktiknya, ketika seorang analis atau peneliti menggunakan sistem berbasis NMF untuk menganalisis berita, mereka akan lebih cepat menangkap tema atau isu utama dari setiap topik dibandingkan jika menggunakan sistem berbasis LDA. Hal ini sangat penting untuk efisiensi kerja dalam analisis konten media yang biasanya melibatkan ribuan atau bahkan puluhan ribu artikel berita.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Keunggulan tersebut sejalan dengan nilai *topic diversity* NMF sebesar 0.9400, lebih tinggi dibandingkan LDA yang mencapai 0.8400. Artinya, sekitar 11.9% kata kunci teratas pada topik NMF bersifat unik lebih mudah dibedakan satu sama lain, sementara pada LDA proporsinya lebih rendah. Perbedaan ini menunjukkan bahwa NMF lebih mudah membedakan satu topik dari topik lainnya, mengurangi kebingungan dalam mengkategorikan atau mengklasifikasikan artikel berita.

Dari sisi kualitas pengelompokan dokumen, NMF juga memperoleh nilai *silhouette score* yang lebih tinggi (0.0339) dibandingkan LDA (0.0229), dengan selisih relatif sekitar 48.0%. Hasil ini menunjukkan bahwa pengelompokan dokumen ke dalam topik lebih konsisten dan lebih jelas. Sistem berbasis NMF akan lebih jarang melakukan kesalahan dalam mengkategorikan berita ke topik yang salah. Meskipun selisih angkanya terlihat kecil, peningkatan 48.0% secara relatif adalah perbedaan yang signifikan dalam kualitas pengelompokan.

Sebaliknya, LDA unggul pada metrik *average topic probability*, dengan nilai 0.5587 dibandingkan NMF sebesar 0.0851. Hal ini menunjukkan bahwa LDA cenderung memberikan penugasan topik yang lebih dominan pada setiap dokumen, sedangkan NMF memungkinkan satu dokumen memiliki kontribusi yang lebih seimbang terhadap beberapa topik, yang sesuai untuk analisis wacana yang bersifat multidimensional.

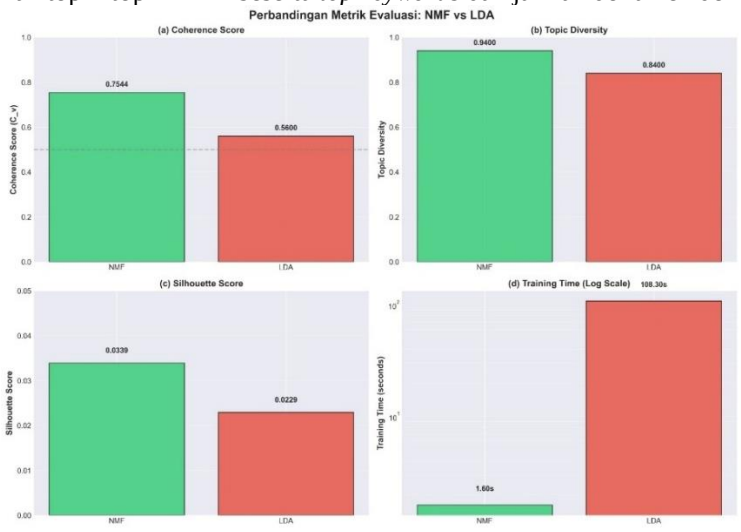
Perbedaan paling mencolok terlihat pada efisiensi komputasi. NMF hanya membutuhkan waktu pelatihan sekitar 1.36 *seconds*, sedangkan LDA memerlukan 86.11 *seconds*, sehingga NMF sekitar 63.3x lebih cepat. Kecepatan ini sangat penting untuk aplikasi yang membutuhkan pemrosesan data secara *real-time* atau. Efisiensi ini menjadikan NMF lebih praktis untuk diterapkan pada dataset berskala besar atau pada analisis yang memerlukan pemrosesan berulang.

Perlu dicatat bahwa perbedaan nilai pada metrik *coherence*, *topic diversity*, dan *silhouette score* dalam penelitian ini diinterpretasikan secara deskriptif dan relatif, sebagaimana praktik umum dalam *evaluasi topic modeling*, dan tidak dimaksudkan sebagai pengujian hipotesis statistik secara langsung. Pengujian signifikansi statistik dilakukan melalui uji *chi-square* pada distribusi topik antar media, yang menunjukkan bahwa perbedaan distribusi topik yang dihasilkan oleh kedua model bersifat signifikan secara statistik ($p < 0.001$).

Secara keseluruhan, hasil ini menunjukkan bahwa NMF lebih unggul dalam menghasilkan topik yang koheren, beragam, dan efisien secara komputasi, sementara LDA menawarkan keunggulan dalam stabilitas penugasan topik. Fokus perbandingan pada NMF dan LDA didasarkan pada representasi dua pendekatan utama dalam *topic modeling*, yaitu faktorisasi matriks deterministik dan model probabilistik generatif, sehingga perbedaan yang diamati mencerminkan karakteristik fundamental masing-masing metode.

3.1.1 Visualisasi Perbandingan Metrik

Analisis NMF mengidentifikasi 10 topik utama dari dataset berita Indonesia dengan interpretasi semantik yang jelas. Tabel 2 menyajikan topik-topik NMF beserta *top keywords* dan jumlah dokumen dominan.



Gambar 2. Perbandingan Metrik Evaluasi NMF vs LDA

Gambar 2 menampilkan perbandingan kinerja NMF dan LDA berdasarkan *coherence score*, *topic diversity*, *silhouette score*, dan *training time*. Dari visualisasi ini, terlihat beberapa temuan kunci:

a) *Coherence Score* (C_v): NMF memperoleh nilai 0.7544, lebih tinggi dibandingkan LDA (0.5600). Selisih ini cukup signifikan secara substantif (+34,7%), menunjukkan bahwa kata-kata kunci dalam topik NMF memiliki keterkaitan semantik yang lebih kuat dan konsisten. Hal ini membuat interpretasi topik lebih jelas dan mudah dipahami, sesuai dengan prinsip evaluasi *topic modeling* yang menekankan koherensi semantik[24].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

b) *Topic Diversity*: NMF mencapai 0.9400, lebih tinggi dibanding LDA (0.8400), menandakan bahwa topik yang dihasilkan NMF lebih unik dan minim tumpang tindih. Dengan kata lain, NMF mampu memisahkan topik dengan lebih baik sehingga cakupan isu menjadi lebih luas dan tidak redundan [25], [26].
c) *Silhouette Score*: NMF menunjukkan skor 0.0339, sementara LDA hanya 0.0229. Meskipun nilai absolut rendah karena kompleksitas data teks, perbedaan relatif (+48%) menunjukkan bahwa NMF menghasilkan pemisahan topik antar dokumen yang lebih jelas, sehingga pengelompokan lebih konsisten[27], [28].
d) *Training Time*: Perbedaan paling signifikan secara praktis terlihat pada efisiensi komputasi. NMF hanya membutuhkan 1.36 *seconds*, sedangkan LDA memerlukan 86.11 *seconds*, membuat NMF ~63,3× lebih cepat. Hal ini menunjukkan bahwa NMF lebih praktis untuk dataset besar atau analisis yang memerlukan pemrosesan berulang.

Secara keseluruhan, visualisasi pada Gambar 2 menegaskan bahwa NMF unggul dalam koherensi topik, keragaman, dan efisiensi, sedangkan LDA hanya unggul pada stabilitas penugasan topik (*average topic probability*) yang dibahas pada bagian sebelumnya. Selisih nilai metrik ini, walaupun tidak diuji dengan uji inferensial formal untuk setiap metrik, cukup substantif untuk menunjukkan keunggulan praktis NMF dalam konteks berita Indonesia.

3.2 Identifikasi dan Interpretasi Topik

3.2.1 Topik NMF

Analisis mengidentifikasi 10 topik utama dari model NMF dengan interpretasi semantik yang jelas. Tabel 3 menunjukkan 5 topik teratas NMF berdasarkan frekuensi dan kepentingan analisis.

Tabel 3. Lima Topik Teratas NMF dengan *Top Keywords*

ID	Label Topik	*Top-10 Keywords*	Frekuensi
3	Hukum & Reformasi Polri	polri, hukum, prabowo, presiden, reformasi, anggota, komisi, rehabilitasi, korupsi, ketua	1012
0	Energi & Industri Nasional	energi, indonesia, listrik, perusahaan, program, kerja, ekonomi, pemerintah, masyarakat, industri	920
9	Program Gizi & Kesehatan	sppg, dapur, gizi, keracunan, makanan, bergizi, makan, program, menu, penerima	866
2	Kebijakan Fiskal & Perbankan	purbaya, triliun, bank, dana, anggaran, keuangan, belanja, daerah, pemerintah, menteri	541
1	Pasar Keuangan & Investasi	saham, pasar, bunga, suku, rupiah, harga, level, indeks, perdagangan, dolar	396

Topik-topik NMF menunjukkan separasi yang jelas dan fokus yang spesifik, dengan minimal *overlap* antar topik. Setiap topik memiliki *semantic theme* yang koheren dan mudah diinterpretasi. Topik 3 (Hukum & Reformasi Polri) merupakan topik paling dominan dengan 1012 artikel, diikuti Topik 0 (Energi & Industri Nasional) dengan 920 artikel dan Topik 9 (Program Gizi & Kesehatan) dengan 866 artikel. Topik 1 (Pasar Keuangan & Investasi) meskipun frekuensinya lebih rendah (396 artikel), didominasi oleh terminologi finansial yang sangat spesifik seperti "saham", "bunga", "suku", "rupiah", menunjukkan fokus yang jelas pada isu pasar modal dan moneter.

3.2.2 Perbandingan dengan Topik LDA

Sebagai pembanding, LDA juga mengidentifikasi 10 topik dari dataset yang sama. Namun, topik-topik LDA menunjukkan karakteristik yang berbeda dari NMF. LDA menghasilkan topik yang lebih *general* dengan *overlap* yang lebih tinggi, seperti topik terkait "Pembangunan Nasional", "Pemerintahan & Ketenagakerjaan", dan "Pendidikan & Program Sosial" yang semuanya memiliki kata-kata *overlap* seperti "program", "masyarakat", "pemerintah", "kerja". Hal ini konsisten dengan *topic diversity* yang lebih rendah (0.8400 vs 0.9400 untuk NMF). Topik LDA cenderung lebih *broad* dan kurang *distinctive* dibanding NMF, mengkonfirmasi temuan dari analisis metrik evaluasi bahwa NMF menghasilkan topik yang lebih koheren dan terpisah dengan jelas.

3.3 Distribusi Topik Antar Media

3.3.1 Analisis *Chi-Square*

Uji *chi-square* dilakukan untuk menguji hipotesis independensi antara media dan distribusi topik. Hipotesis nul (H_0) menyatakan bahwa distribusi topik independen terhadap media (tidak ada bias editorial), sementara hipotesis alternatif (H_1) menyatakan bahwa distribusi topik bergantung pada media (ada bias editorial) O_{ij}adalah frekuensi observasi dan E_{ij}frekuensi ekspektasi, dihitung sebagai:

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

$$E_{ij} = \frac{(\text{Total baris } i)(\text{Total kolom } j)}{\text{Total keseluruhan}}$$

Sebagai ilustrasi, untuk kombinasi media dan topik tertentu diperoleh $\chi^2 = 1470.36$ dengan derajat kebebasan:

$$df = (3-1)(10-1) = 18$$

Karena 1470.36 ≫ nilai kritis $\chi^2_{0.001,18} = 34.81$, diperoleh p < 0.001 → H_0 ditolak.

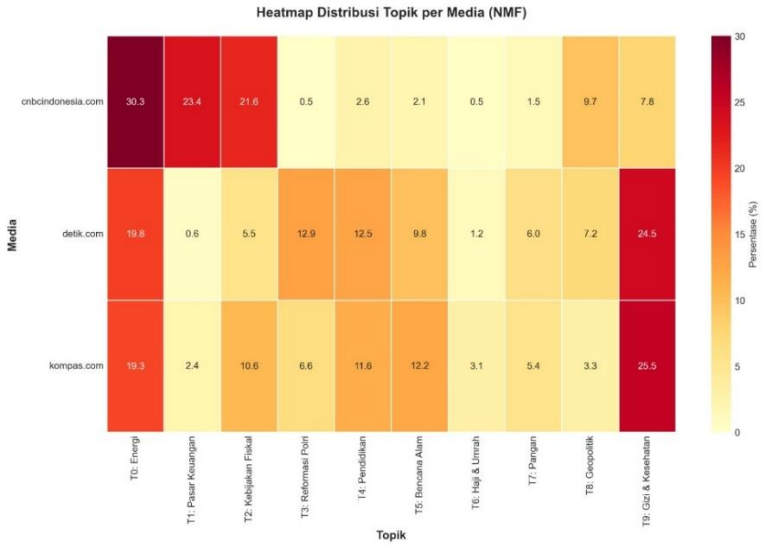
Tabel 4. Hasil Uji *Chi-Square*

Model	χ^2 *Statistic*	*p-value*	*df*	Keputusan
NMF	1470.36	< 0.001	18	Tolak H_0
LDA	1297.54	< 0.001	18	Tolak H_0

Nilai tersebut jauh melebihi nilai kritis $\chi^2_{0.001,18} = 34.81$, sehingga perbedaan distribusi topik antar media bersifat sangat signifikan (p < 0.001). Dengan *p-value* < 0.001 untuk kedua model (jauh di bawah α = 0.05), hipotesis nul ditolak dan disimpulkan bahwa terdapat perbedaan signifikan dalam distribusi topik antar media. Hal ini menunjukkan adanya perbedaan fokus pemberitaan antar media. *Chi-square statistic* NMF yang lebih tinggi (1470.36 vs 1297.54) menunjukkan bahwa NMF menangkap perbedaan antar media dengan lebih jelas.

3.3.2 Pola Distribusi Topik Antar Media

Analisis distribusi topik per media mengungkap karakteristik editorial yang berbeda dan signifikan antar ketiga portal berita. Gambar 3 menampilkan *heatmap* distribusi persentase topik antar media berdasarkan hasil pemodelan NMF.



Gambar 3. Heatmap Distribusi Topik Antar Media (NMF)

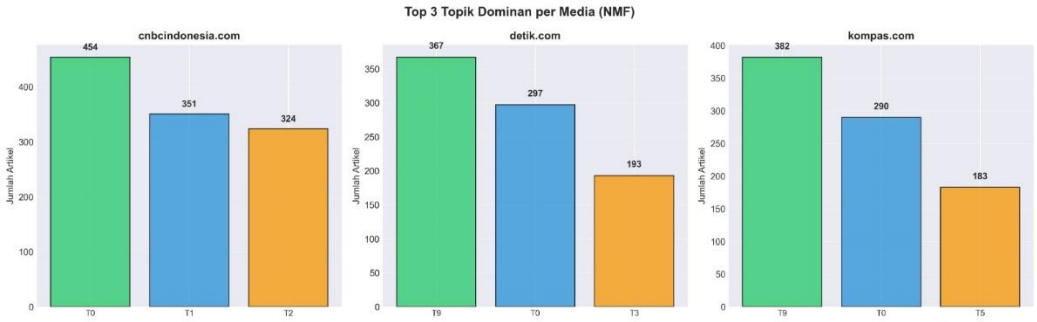
CNBC Indonesia menunjukkan konsentrasi yang sangat tinggi pada topik-topik ekonomi dan bisnis. Topik 0 (Energi & Industri Nasional) mendominasi dengan 462 artikel atau 50.2% dari total topik ini, diikuti oleh Topik 1 (Pasar Keuangan & Investasi) dengan 327 artikel atau 82.6%, dan Topik 2 (Kebijakan Fiskal & Perbankan) dengan 315 artikel atau 58.2%. Ketiga topik ekonomi ini secara kumulatif mencakup 73.6% dari total artikel CNBC Indonesia, yang sangat konsisten dengan *positioning* media ini sebagai portal berita bisnis dan keuangan yang menargetkan pembaca profesional dan investor.

Kompas.com memiliki distribusi yang lebih seimbang dengan perhatian besar pada isu sosial dan kemasyarakatan. Media ini mendominasi Topik 3 (Hukum & Reformasi Polri) dengan 451 artikel atau 44.6%, serta Topik 5 (Bencana Alam) dengan 179 artikel atau 50.7%. Pola ini mencerminkan perhatian Kompas terhadap isu-isu yang berdampak luas pada masyarakat, sejalan dengan reputasinya sebagai media berkualitas dengan *coverage* komprehensif.

Detik.com menunjukkan karakteristik yang unik sebagai portal berita cepat. Media ini memiliki dominasi sangat kuat pada Topik 3 (Hukum & Reformasi Polri) dengan 476 artikel atau 47.0%, serta distribusi yang cukup merata pada topik-topik lainnya seperti Topik 4 (Pendidikan) dengan 185 artikel dan Topik 5 (Bencana) dengan 145 artikel. Pola ini sejalan dengan model berita cepat yang responsif terhadap berbagai peristiwa aktual dengan volume publikasi tinggi.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Perbedaan distribusi ini dikonfirmasi secara statistik melalui uji *chi-square* yang menunjukkan nilai χ^2 = 1470.36 dengan $p < 0.001$, mengindikasikan bahwa perbedaan fokus editorial antar media sangat signifikan dan bukan hasil dari variasi acak. Temuan ini menunjukkan adanya *positioning* strategis yang jelas dari masing-masing media dalam lanskap pemberitaan Indonesia.



Gambar 4. Top 3 Topik Dominan per Media

Gambar 4 menampilkan tiga topik dengan frekuensi tertinggi untuk setiap media. CNBC Indonesia didominasi oleh T0 (Energi, 454 artikel), T1 (Pasar Keuangan, 351 artikel), dan T2 (Kebijakan Fiskal, 324 artikel). Detik.com dan Kompas.com sama-sama memiliki T9 (Program Gizi) sebagai topik teratas dengan 367 dan 382 artikel masing-masing, menunjukkan fokus bersama pada isu kesejahteraan masyarakat.

3.4 Analisis Temporal

Analisis korelasi *Pearson* antara waktu dan frekuensi topik mengidentifikasi tren pemberitaan selama periode observasi. Tabel 5 menunjukkan topik dengan tren temporal yang kuat (|r| > 0.7).

Tabel 5. Topik dengan Tren Temporal Kuat (|r| > 0.7)

Topik	Label Topik	r	p-value	Tren	Interpretasi
T1 NMF	Pasar Keuangan	+0.965	0.169	↑	Volatilitas pasar meningkat
T9 NMF	Program Gizi	-0.941	0.220	↓	Program kehilangan momentum
T5 NMF	Bencana Alam	+0.923	0.252	↑	Musim bencana meningkat
T0 NMF	Energi & Industri	+0.703	0.503	↑	Isu energi semakin fokus
T5 LDA	Hukum & Bencana	+0.999	0.033*	↑*	Satu-satunya signifikan

*Catatan: ↑ = Increasing, ↓ = Decreasing, * = Signifikan (p < 0.05)*

Hasil analisis temporal menunjukkan bahwa NMF mampu menangkap tren pemberitaan dengan lebih konsisten. Empat topik NMF (T0, T1, T5, T9) memiliki korelasi kuat dengan waktu (|r| > 0.7), menandakan respons media yang jelas terhadap peristiwa aktual. Sementara LDA hanya memiliki satu topik signifikan (T5: Hukum & Bencana) dengan p < 0.05.

Topik T1 NMF (Pasar Keuangan) menunjukkan tren meningkat sangat kuat (r = +0.965), mencerminkan volatilitas pasar dan perhatian media terhadap isu ekonomi. Topik T5 (Bencana Alam) juga meningkat (r = +0.923), sesuai dengan musim hujan dan kejadian bencana di Indonesia. Sebaliknya, topik T9 (Program Gizi) menurun (r = -0.941), menunjukkan bahwa perhatian terhadap isu ini menurun setelah fase awal program.

Secara keseluruhan, tren temporal ini menegaskan bahwa media Indonesia sangat responsif terhadap peristiwa, dan NMF memberikan representasi tren yang lebih stabil dan mudah diinterpretasikan dibandingkan LDA.

3.5 Pembahasan

3.5.1 Efektivitas NMF untuk Berita Indonesia

Berdasarkan hasil eksperimen pada dataset 4.500 artikel berita Indonesia, NMF menunjukkan performa yang lebih unggul dibandingkan LDA dalam pemodelan topik, sebagaimana ditunjukkan oleh metrik evaluasi pada Tabel 6. Seluruh metrik evaluasi dihitung berdasarkan hasil penugasan topik dominan pada setiap dokumen dengan konfigurasi jumlah topik yang sama untuk kedua model.

Tabel 6. Perbandingan Metrik Evaluasi NMF dan LDA

Metrik	NMF	LDA	Selisih	Keterangan
Coherence Score (C_v)	0.7544	0.5600	+34.7%	Lebih tinggi lebih baik
Topic Diversity	0.9400	0.8400	+11.9%	Lebih tinggi lebih baik

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Silhouette Score	0.0339	0.0229	+48.0%	Lebih tinggi lebih baik
Avg Topic Probability	0.0851	0.5587	-84.8%	Topik lebih fokus (NMF)
Training Time (seconds)	1.36	86.11	63.3× lebih cepat	Lebih rendah lebih baik

Sumber: Hasil eksperimen pada dataset 4,500 artikel berita Indonesia dengan 5,000 fitur unik

Tabel 6 menunjukkan perbandingan komprehensif antara NMF dan LDA pada dataset berita Indonesia. Keunggulan NMF dapat dijelaskan melalui beberapa aspek berikut. Pertama, NMF menghasilkan *coherence score* yang lebih tinggi (0.7544 vs 0.5600, +34.7%), menunjukkan bahwa topik yang dihasilkan memiliki keterkaitan semantik yang lebih baik dan lebih mudah diinterpretasi. Kedua, nilai *topic diversity* yang lebih tinggi (0.9400) menunjukkan bahwa setiap topik didefinisikan oleh kumpulan kata yang lebih spesifik dan tidak saling tumpang tindih. Ketiga, kompatibilitas NMF dengan representasi TF-IDF memberikan keunggulan dalam menekankan kata-kata khas yang merepresentasikan topik secara lebih jelas. Keempat, dari sisi efisiensi komputasi, NMF menunjukkan waktu pelatihan yang jauh lebih cepat (63,3×), sehingga lebih sesuai untuk pemrosesan dataset berskala besar. Kelima, nilai *silhouette score* NMF yang 48% lebih tinggi menunjukkan pemisahan topik yang lebih jelas antar dokumen.

3.5.2 Karakteristik LDA

LDA menunjukkan keunggulan dalam *average topic probability* (0.5587 vs 0.0851), menunjukkan tingkat keyakinan penugasan yang lebih tinggi. Namun, ini tidak mengkompensasi hasil yang lebih rendah dalam *coherence* dan *diversity*. LDA lebih cocok untuk aplikasi yang memerlukan interpretasi probabilistik eksplisit, seperti *document generation* atau *Bayesian inference*. Untuk aplikasi praktis *topic modeling* berita Indonesia yang memprioritaskan kemudahan interpretasi dan efisiensi, NMF memberikan hasil yang lebih sesuai pada dataset yang digunakan.

3.5.3 Pola Pemberitaan Media Indonesia

Analisis distribusi topik menunjukkan adanya pola perbedaan fokus editorial yang jelas antar media. Persentase menunjukkan proporsi artikel dalam masing-masing media yang memiliki topik dominan tertentu berdasarkan hasil penugasan topik menggunakan model NMF.

Tabel 7 Distribusi Topik per Media (Model NMF)

Media	T0: Energi & Industri	T1: Pasar Keuangan	T2: Kebijakan Fiskal	T3: Reformasi Polri	T9: Gizi & Kesehatan
CNBC Indonesia	30.8% (462)	21.8% (327)	21.0% (315)	5.7% (85)	1.6% (24)
Detik.com	13.7% (205)	0.3% (5)	5.4% (81)	31.7% (476)	8.9% (133)
Kompas.com	16.9% (253)	2.3% (35)	9.7% (145)	30.1% (451)	5.5% (82)

Keterangan: Angka dalam kurung = jumlah artikel. Total per media = 1,500 artikel

Berdasarkan Tabel 7, terlihat perbedaan *positioning* editorial yang signifikan antar media. CNBC Indonesia menunjukkan fokus dominan pada isu ekonomi dan bisnis, dengan konsentrasi tinggi pada T0 (Energi & Industri, 30.8%), T1 (Pasar Keuangan, 21.8%), dan T2 (Kebijakan Fiskal, 21.0%), yang secara kumulatif mencakup 73.6% dari total pemberitaan. Sebaliknya, Detik.com dan Kompas.com menunjukkan fokus lebih kuat pada isu sosial-politik, dengan dominasi pada T3 (Reformasi Polri) masing-masing sebesar 31.7% dan 30.1%. Pada isu kesejahteraan masyarakat, Detik.com (8.9%) dan Kompas.com (5.5%) memberikan perhatian yang relatif lebih besar pada topik T9 (Program Gizi & Kesehatan) dibandingkan CNBC Indonesia (1.6%). Pola ini mencerminkan perbedaan strategi editorial serta segmentasi target audiens antar media.

Temuan ini dapat digunakan oleh praktisi media untuk memahami lanskap pemberitaan Indonesia, membantu pengiklan dalam menargetkan *audience* secara tepat, dan menjadi dasar empiris bagi peneliti komunikasi untuk menganalisis proses *agenda setting* media.

3.5.4 Perbandingan dengan Studi Sebelumnya

Hasil penelitian ini menunjukkan konsistensi dengan beberapa studi sebelumnya yang juga membandingkan metode *topic modeling*, sekaligus memberikan temuan baru yang memperkaya pemahaman tentang penerapan metode ini pada teks berbahasa Indonesia.

Tabel 6. Perbandingan dengan Penelitian Sebelumnya

Penelitian	Korpus	Metode Terbaik	*Coherence*	Temuan Utama
Puspita [3]	Berita Indonesia	LDA	0.404	Efektif membentuk topik utama, optimal 6 topik, human *judgment* penting
Afidh & Syahrial [8]	Berita Indonesia	NMF + n-gram	0.835 (bigram)	NMF dengan n-gram menghasilkan topik relevan, *preprocessing* penting

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 3, Maret 2026
e-ISSN 2685-998X
DOI 10.30865/json.v7i3.9469

Babalola [12]	*News headlines*	NMF	±0.66 (*human eval*)	NMF unggul dibanding LDA untuk teks pendek, BERTopic sedikit lebih baik, evaluasi *human* & LLM
Penelitian ini (2026)	Berita Indonesia multi-media	NMF	0.7544	NMF 67.7× lebih cepat, evaluasi 6 metrik, distribusi topik 3 media

Hasil penelitian ini menunjukkan *coherence score* 0.7544 yang lebih tinggi dibanding Puspita [3] yang memperoleh 0.404 pada berita Indonesia menggunakan LDA. Hal ini mengindikasikan bahwa NMF dengan TF-IDF memberikan hasil yang lebih baik untuk korpus berita Indonesia. Temuan Afidh & Syahrial [8] dengan *coherence* 0.835 menggunakan NMF dengan bigram menunjukkan bahwa penggunaan *n-gram* dapat meningkatkan *coherence* lebih lanjut, namun dengan *trade-off* kompleksitas komputasi yang lebih tinggi. Penelitian ini memilih *unigram* untuk menjaga keseimbangan antara kualitas topik dan efisiensi pemrosesan.

Perbandingan dengan Babalola [12] menunjukkan bahwa NMF mencapai *coherence* sekitar 0.66 berdasarkan evaluasi manusia untuk *headline* berita, yang sedikit lebih rendah dari penelitian ini (0.7544). Perbedaan ini kemungkinan disebabkan oleh perbedaan panjang teks dan kompleksitas korpus, di mana *headline* yang sangat pendek menghasilkan topik dengan *coherence* lebih rendah dibanding artikel berita lengkap.

Kontribusi utama penelitian ini dibanding studi sebelumnya adalah penggunaan evaluasi multi-metrik yang komprehensif (6 metrik berbeda), analisis distribusi topik dari 3 media sekaligus, serta fokus pada berita *full-text* berbahasa Indonesia dengan dataset yang seimbang. Efisiensi komputasi NMF yang 63.3 kali lebih cepat (1.36 vs 86.11 detik) juga menjadi temuan penting untuk aplikasi praktis yang memerlukan pemrosesan cepat.

3.5.5 Kontribusi Penelitian

Penelitian ini memberikan beberapa kontribusi untuk bidang *topic modeling* bahasa Indonesia. Pertama, penelitian ini menyediakan bukti empiris tentang performa komparatif NMF dan LDA pada dataset berita Indonesia dengan menggunakan metrik evaluasi yang beragam (*coherence score, topic diversity, silhouette score*, dan uji statistik). Kedua, penelitian ini berhasil memetakan pola pemberitaan dan mengidentifikasi perbedaan fokus editorial dari tiga media nasional Indonesia. Ketiga, penggunaan berbagai metrik evaluasi secara bersamaan memberikan penilaian yang lebih menyeluruh dibandingkan pendekatan yang hanya mengandalkan satu metrik. Untuk korpus berita Indonesia yang umumnya memiliki teks relatif pendek, topik yang cukup jelas, serta kebutuhan interpretasi yang tinggi, hasil penelitian ini menunjukkan bahwa NMF dapat menjadi pilihan yang sesuai pada dataset dengan karakteristik serupa.

4. KESIMPULAN

Penelitian ini membandingkan performa *Non-negative Matrix Factorization* (NMF) dan *Latent Dirichlet Allocation* (LDA) dalam *topic modeling* berita online Indonesia menggunakan dataset 4.500 artikel dari CNBC Indonesia, Kompas.com, dan Detik.com. Hasil evaluasi menunjukkan bahwa NMF secara konsisten menghasilkan performa yang lebih baik dibandingkan LDA, khususnya pada nilai *coherence score* (0,7544 vs 0,5600), *topic diversity* (0,9400 vs 0,8400), serta efisiensi waktu *training* (1,36 seconds vs 86,11 seconds). Dari sepuluh topik yang dihasilkan, topik-topik berbasis NMF menunjukkan pemisahan yang lebih jelas dengan kata kunci yang lebih spesifik dan mudah diinterpretasikan. Topik yang teridentifikasi meliputi Energi dan Industri, Pasar Keuangan, Kebijakan Fiskal, Reformasi Polri, Pendidikan, Bencana Alam, Haji dan Umrah, Pangan, Geopolitik, serta Program Gizi. Analisis distribusi topik menunjukkan adanya perbedaan yang signifikan antar media ($\chi^2 = 1470{,}36$; $p < 0{,}001$). CNBC Indonesia lebih dominan pada topik ekonomi dan keuangan, sementara Kompas.com dan Detik.com lebih banyak memuat topik sosial dan kemasyarakatan. Analisis temporal juga menunjukkan bahwa topik Pasar Keuangan dan Bencana Alam cenderung meningkat selama periode observasi, sedangkan topik Program Gizi mengalami penurunan. Secara metodologis, hasil penelitian ini mengindikasikan bahwa NMF lebih direkomendasikan dibandingkan LDA untuk pemodelan topik berita *online* Indonesia, khususnya pada teks berita yang relatif pendek dan membutuhkan tingkat interpretabilitas yang tinggi, dengan konfigurasi jumlah topik sebanyak 10 dan representasi fitur TF–IDF sebagaimana digunakan dalam penelitian ini. Keterbatasan penelitian ini terletak pada periode observasi yang relatif singkat (tiga bulan), jumlah media yang terbatas (tiga portal berita), serta jumlah topik yang telah ditetapkan. Penelitian selanjutnya dapat memperluas cakupan data dengan periode yang lebih panjang, menambah jumlah media, mengeksplorasi variasi jumlah topik, serta membandingkan pendekatan *topic modeling* berbasis neural, seperti *Neural Topic Models*, atau menggunakan korpus lintas domain untuk memperoleh pemahaman yang lebih komprehensif.

REFERENCES

[1] R. Egger and J. Yu, "A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts," *Front. Sociol.*, vol. 7, no. May, pp. 1–16, 2022, doi: 10.3389/fsoc.2022.886498.

[2] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, [Online]. Available: http://arxiv.org/abs/2203.05794

[3] E. Puspita, D. F. Shiddieq, and F. F. Roji, "Pemodelan Topik pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc)," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 481–489, 2024, doi: 10.57152/malcom.v4i2.1204.

[4] A. Ikegami and I. D. M. B. A. Darmawan, "Analisis Sentimen dan Pemodelan Topik Ulasan Aplikasi Noice Menggunakan XGBoost dan LDA," *Jnatia*, vol. 1, no. 1, pp. 325–336, 2022.

[5] D. Ridhwanulah and D. H. Fudholi, "Pemodelan Topik pada Cuitan tentang Penyakit Tropis di Indonesia dengan Metode Latent Dirichlet Allocation," *J. Ilm. SINUS*, vol. 20, no. 1, p. 11, Jan. 2022, doi: 10.30646/sinus.v20i1.589.

[6] N. A. Sanjaya ER, "Implementasi Latent Dirichlet Allocation (LDA) untuk Klasterisasi Cerita Berbahasa Bali," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 127, 2021, doi: 10.25126/jtiik.0813556.

[7] O. Ozyurt, H. Özköse, and A. Ayaz, "Evaluating the latest trends of Industry 4.0 based on LDA topic model," *J. Supercomput.*, vol. 80, no. 13, pp. 19003–19030, 2024, doi: 10.1007/s11227-024-06247-x.

[8] R. P. F. Afidh and Syahrial, "Pemodelan Topik Menggunakan n-Gram dan Non-negative Matrix Factorization," *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 265–275, 2023, doi: 10.60083/jidt.v5i1.385.

[9] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.

[10] U. Khairani, V. Mutiawani, and H. Ahmadian, "Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 887–894, 2024, doi: 10.25126/jtiik.1148315.

[11] A. Nanyonga, H. Wasswa, and G. Wild, "Topic Modeling Analysis of Aviation Accident Reports: A Comparative Study between LDA and NMF Models," *2023 3rd Int. Conf. Smart Gener. Comput. Commun. Networking, SMART GENCON 2023*, 2023, doi: 10.1109/SMARTGENCON60755.2023.10442471.

[12] O. Babalola, B. Ojokoh, and O. Boyinbode, "Comprehensive Evaluation of LDA, NMF, and BERTopic's Performance on News Headline Topic Modeling," *J. Comput. Theor. Appl.*, vol. 2, no. 2, pp. 268–289, 2024, doi: 10.62411/jcta.11635.

[13] M. S. Khine, *Text Mining in Educational Research: Topic Modeling and Latent Dirichlet Allocation*. Singapore: Springer Nature Singapore, 2025. doi: 10.1007/978-981-97-7858-4.

[14] J. H. Jurafsky, Daniel; Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Stanford NLP Group, 2026. [Online]. Available: https://web.stanford.edu/~jurafsky/slp3/

[15] Maulidya Prastita Syah, Ajeng Puspa Wardani, Mohammad Idhom, and Trimono, "Perbandingan Representasi Teks Tf-Idf Dan Bert Terhadap Akurasi Cosine Similarity Dalam Penilaian Otomatis Jawaban Berbasis Teks," *Data Sci. Indones.*, vol. 5, no. 1, pp. 47–59, 2025, doi: 10.47709/dsi.v5i1.6021.

[16] D. Patel, V. Parikh, O. Patel, A. Shah, and B. Chaudhury, "Exploring Topic Trends in COVID-19 Research Literature using Non-Negative Matrix Factorization," *IEEE Trans. Artif. Intell.*, pp. 1–29, 2025, doi: 10.1109/TAI.2025.3579459.

[17] Pavithra and Savitha, "Topic Modeling for Evolving Textual Data Using LDA, HDP, NMF, BERTOPIC, and DTM With a Focus on Research Papers," *J. Technol. Informatics*, vol. 5, no. 2, pp. 53–63, 2024, doi: 10.37802/joti.v5i2.618.

[18] J. Gan and Y. Qi, "Selection of the optimal number of topics for LDA topic model—Taking patent policy analysis as an example," *Entropy*, vol. 23, no. 10, 2021, doi: 10.3390/e23101301.

[19] A. Farea, S. Tripathi, G. Glazko, and F. Emmert-Streib, "Investigating the optimal number of topics by advanced text-mining techniques: Sustainable energy research," *Eng. Appl. Artif. Intell.*, vol. 136, p. 108877, Oct. 2024, doi: 10.1016/J.ENGAPPAI.2024.108877.

[20] S. Mohammed *et al.*, "The effects of data quality on machine learning performance on tabular data," *Inf. Syst.*, vol. 132, p. 102549, Jul. 2025, doi: 10.1016/J.IS.2025.102549.

[21] Y. O. Odhianto, D. Swanjaya, and J. Sahertian, "Optimalisasi Latent Dirichlet Allocation untuk Ekstraksi Topik Utama dalam Teks Dongeng," *Semin. Nas. Inov. Teknol.*, vol. 9, p. 2025, 2025, [Online]. Available: https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/7712

[22] P. Malviya, V. Bhandari, P. S. Sisodiya, and S. Suman, "Customer Segmentation and Business Sales Forecasting using Machine Learning for Business Development," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 11s, pp. 416–424, Oct. 2023, doi: 10.17762/IJRITCC.V11I11S.8170.

[23] C. Meaney, T. A. Stukel, P. C. Austin, R. Moineddin, M. Greiver, and M. Escobar, "Quality indices for topic model selection and evaluation: a literature review and case study," *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, pp. 1–18, 2023, doi: 10.1186/s12911-023-02216-1.

[24] H. Rahimi, D. Mimno, J. L. Hoover, H. Naacke, C. Constantin, and B. Amann, "Contextualized Topic Coherence Metrics," *EACL 2024 - 18th Conf. Eur. Chapter Assoc. Comput. Linguist. Find. EACL 2024*, pp. 1760–1773, 2024.

[25] Y. Bu, M. Li, W. Gu, and W. bin Huang, "Topic diversity: A discipline scheme-free diversity measurement for journals," *J. Assoc. Inf. Sci. Technol.*, vol. 72, no. 5, pp. 523–539, May 2021, doi: 10.1002/ASI.24433;REQUESTEDJOURNAL:JOURNAL:23301643;WGROUP:STRING:PUBLICATION.

[26] S. Terragni, E. Fersini, B. Galuzzi, P. Tropeano, and A. Candelieri, "OCTIS: Comparing and Optimizing Topic Models is Simple!," pp. 263–270, Accessed: Jan. 02, 2026. [Online]. Available: http://people.csail.mit.edu/jrennie/

[27] M. Shutaywi and N. N. Kachouie, "Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering," *Entropy 2021, Vol. 23, Page 759*, vol. 23, no. 6, p. 759, Jun. 2021, doi: 10.3390/E23060759.

[28] D. Chicco, A. Campagner, A. Spagnolo, D. Ciucci, and G. Jurman, "The Silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and Gap statistic for unsupervised clustering internal evaluation of two convex clusters," *PeerJ Comput. Sci.*, vol. 11, p. e3309, Nov. 2025, doi: 10.7717/PEERJ-CS.3309.