

## Analisis Sentimen Pemberitaan kompas.com Terhadap Program Merdeka Belajar Dengan Metode Clustering K-Means

Dimas Alfiyanto<sup>\*1</sup>, Ahmad Nur Fajri<sup>2</sup>, Erwin Liando<sup>3</sup>, Farhan Dwi Prasetio<sup>4</sup>, Riza Fahlapi<sup>5</sup>, Fachri Amsury<sup>6</sup>  
<sup>1,2,3,4,5,6</sup> Universitas Bina Sarana Informatika; Jakarta Pusat, Indonesia

e-mail: [\\*117230724@bsi.ac.id](mailto:*117230724@bsi.ac.id), [217230079@bsi.ac.id](mailto:217230079@bsi.ac.id), [317230857@bsi.ac.id](mailto:317230857@bsi.ac.id) [4117230827@bsi.ac.id](mailto:4117230827@bsi.ac.id),  
[5riza.rzf@bsi.ac.id](mailto:5riza.rzf@bsi.ac.id), [6fachri.fcy@bsi.ac.id](mailto:6fachri.fcy@bsi.ac.id)

**Abstract –** This study aims to analyze the sentiment of Kompas.com news coverage regarding the Merdeka Belajar Program through the application of the K-Means Clustering algorithm. A text mining approach was applied to 253 news articles published throughout the period of January 2023 to December 2024. The methodological stages included text preprocessing, TF-IDF feature extraction, and clustering into three sentiment groups. The sentiment distribution showed a dominance of positive tone with 69.6% of the articles, compared to 28.5% negative, and 2.0% neutral. Keyword analysis revealed that the pro cluster focused on educational and cultural aspects, while the contra cluster highlighted implementation challenges experienced by schools and students. Although a Silhouette Score value of 0.018 was obtained, indicating limitations in grouping quality due to the complexity of journalistic language, this finding consistently confirmed the tendency for positive coverage from Kompas.com. Methodologically, this study contributes to the development of a sentiment analysis framework for Indonesian-language texts.

**Keywords —** Sentiment Analysis, Freedom to Learn Program, K-Means Clustering, Kompas.com, Text Mining, Indonesian Education.

**Abstrak –** Penelitian ini bertujuan menganalisis sentimen pemberitaan Kompas.com terkait Program Merdeka Belajar melalui penerapan K-Means Clustering. Pendekatan text mining diaplikasikan pada 253 artikel berita yang terbit sepanjang Januari 2023 hingga Desember 2024. Tahapan metodologi mencakup prapemrosesan teks, ekstraksi fitur TF-IDF, dan clustering ke dalam tiga kelompok sentimen. Distribusi sentimen menunjukkan dominasi nada positif dengan 69,6% artikel, berbanding 28,5% negatif, dan 2,0% netral. Analisis kata kunci klaster pro berpusat pada aspek didik dan budaya, sementara klaster kontra menggarisbawahi tantangan implementasi yang dialami sekolah dan siswa. Meskipun nilai Silhouette Score sebesar 0,018 mengindikasikan adanya batasan pengelompokan akibat kompleksitas bahasa jurnalistik, temuan ini secara konsisten mengonfirmasi kecenderungan pemberitaan positif Kompas.com. Secara metodologis, studi ini memberikan kontribusi dalam pengembangan kerangka analisis sentimen untuk teks-teks dalam Bahasa Indonesia.

**Kata Kunci —** Analisis Sentimen, Program Merdeka Belajar, K-Means Clustering, Kompas.com, Penambangan Teks, Pendidikan Indonesia.

### I. PENDAHULUAN

Pendidikan merupakan pilar fundamental dalam pembangunan sumber daya manusia suatu bangsa. Untuk menjawab tantangan global dan kebutuhan zaman, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbud Ristek) telah menginisiasi Program Merdeka Belajar. Program ini menjadi terobosan penting yang bertujuan mentransformasi sistem pendidikan nasional dengan menekankan fleksibilitas, kreativitas, dan otonomi yang lebih besar bagi institusi pendidikan, guru, dan peserta didik.

Tingginya urgensi program ini tercermin jelas dalam diskursus publik dan pemberitaan media massa. Media online terkemuka seperti Kompas.com memiliki pengaruh signifikan dalam membentuk persepsi publik terhadap kebijakan ini. Data menunjukkan bahwa isu pendidikan, khususnya Program Merdeka Belajar, mendominasi liputan pers nasional, dengan total lebih dari 350 artikel yang berhasil diakses pada periode studi. Pemberitaan ini mengandung nuansa sentimen yang penting untuk dianalisis, sebab media mainstream menjalankan fungsi gatekeeping dan framing dalam mengonstruksi realitas Program Merdeka Belajar di mata pembaca.

Meskipun penelitian sebelumnya telah mengkaji analisis sentimen isu pendidikan di media sosial, studi-studi tersebut membatasi fokus pada data yang tidak terstruktur dan tidak mewakili pandangan media institusional. Oleh karena itu, penelitian ini secara eksplisit mengisi research gap dengan fokus pada analisis sentimen pemberitaan media mainstream berkapasitas tinggi, yaitu Kompas.com, selama periode Januari 2023 hingga Desember 2024. Untuk mencapai tujuan tersebut, digunakan pendekatan text mining dan algoritma K-Means Clustering yang efektif dalam mengelompokkan data teks berita yang besar dan kompleks secara tak-terawasi (unsupervised).

Tujuan utama penelitian ini adalah menganalisis polaritas sentimen dan mengidentifikasi pola framing dalam pemberitaan Kompas.com terhadap Program Merdeka Belajar. Kontribusi ilmiah yang dihasilkan meliputi: Pengujian efektivitas algoritma clustering K-Means dengan fitur TF-IDF pada data teks berita berbahasa Indonesia yang memiliki kompleksitas tinggi. Dan Penyediaan wawasan empiris mengenai polaritas sentimen dan pola framing media, yang dapat digunakan sebagai referensi evaluasi komunikasi kebijakan dan literasi media.

## II. PENELITIAN YANG TERKAIT

### A. Konsep Dasar Program Merdeka Belajar

Program Merdeka Belajar merupakan inisiatif utama Kemendikbud Ristek sejak 2020 yang bertujuan mentransformasi pendidikan nasional [3]. Fokus program ini adalah memberikan fleksibilitas, otonomi, dan kemerdekaan bagi institusi pendidikan dan guru untuk berinovasi, sejalan dengan kebutuhan Society 5.0 [14]. Pilar utamanya mencakup pengembangan kurikulum adaptif, peningkatan kualitas guru, pembelajaran berpusat pada siswa, dan kolaborasi dengan industri [5].

### B. Landasan Teori Analisis Sentimen dan Text Mining

Analisis Analisis Sentimen adalah bidang ilmu yang mengukur pendapat, perasaan, dan sikap publik terhadap suatu entitas atau topik [7]. Pendekatan dapat dilakukan melalui metode berbasis lexicon, machine learning (supervised), atau hybrid [8]. Dalam penerapannya, Text Mining berperan sebagai pendekatan untuk mengidentifikasi sentimen dan topik dengan mengubah data teks mentah menjadi fitur yang dapat diolah. Pra-pemrosesan data merupakan tahap krusial dalam text mining [9].

### C. Metode Clustering K-Means dan Fitur TF-IDF

K-Means Clustering merupakan algoritma pembelajaran tanpa pengawasan (*unsupervised*) yang efisien dalam mengelompokkan data berdasarkan kesamaan karakteristik, cocok untuk menganalisis data teks yang tidak berlabel [10]. Pengelompokan ini dilakukan dengan meminimalkan Jarak Euclidean antar titik data dan centroid kluster.

$$D(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

Untuk merepresentasikan data teks, digunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency) sebagai metode ekstraksi fitur. TF-IDF menentukan pembobotan kata dengan menghitung perkalian frekuensi istilah (TF) dalam dokumen dan kelangkaannya di seluruh korpus (IDF) [9] :

$$TF - IDF = TF(t, d) \times IDF(t)$$

### D. Peran Media Mainstream dan Framing

Media memiliki peran signifikan dalam mengonstruksi realitas melalui proses framing, yaitu menyeleksi dan menonjolkan aspek tertentu dari isu untuk mempromosikan pandangan tertentu [11].

Kompas.com, sebagai media mainstream, memiliki kecenderungan framing yang berimbang dan moderat, berfokus pada fakta dan solusi konstruktif [13]. Pemilihan framing ini sangat menentukan sentimen yang diterima oleh publik [12].

### E. Penelitian Terdahulu dan Posisi Penelitian

Tinjauan Tinjauan literatur menunjukkan bahwa studi sentimen terkait Merdeka Belajar cenderung berfokus pada data media sosial dengan metode supervised [1]. sementara aplikasi clustering pada artikel berita online sering kali tidak berfokus pada polaritas kebijakan spesifik [2]. Selain itu, mekanisme framing media online terhadap kebijakan publik telah diteliti diteliti [12].Kesenjangan utama (research gap) yang diisi oleh studi ini adalah belum adanya analisis sentimen secara unsupervised menggunakan K-Means Clustering pada data teks jurnalistik yang diterbitkan oleh media mainstream institusional seperti Kompas.com. Penelitian ini menggabungkan analisis clustering dengan interpretasi framing untuk memberikan wawasan yang komprehensif mengenai konstruksi sentimen media terhadap Program Merdeka Belajar.

## III. METODE PENELITIAN

### A. Workflow Penelitian

Workflow penelitian ini terdiri dari enam tahap utama seperti yang ditunjukkan pada Gambar 1. Tahapan dimulai dari pengumpulan data, prapemrosesan teks, ekstraksi fitur, klasterisasi K-Means, analisis sentimen dan interpretasi, hingga visualisasi hasil.



Gambar 1 Workflow Penelitian

### B. Pengumpulan Data

Tahap awal penelitian adalah pengumpulan 253 artikel berita online yang bersumber dari portal berita Kompas.com dengan fokus isu Program Merdeka Belajar. Akuisisi data dilakukan secara otomatis menggunakan teknik web scraping dengan library Python Scrapy selama periode Januari 2023 hingga Desember 2024.

Untuk memastikan korpus data yang relevan dan berkualitas, kriteria seleksi diterapkan:

- Kriteria Inklusi: Artikel harus membahas Program Merdeka Belajar (termasuk Kampus Merdeka atau MBKM) sebagai konten utama; dipublikasikan dalam periode studi; merupakan jenis hard news, feature news, atau reportase khusus pendidikan; serta memiliki konten minimal 300 kata.
- Kriteria Eksklusi: Artikel dieliminasi jika berupa konten tidak relevan (opini/editorial, iklan/sponsorship), news briefs (kurang dari 300 kata), duplikasi konten (repost), atau memiliki masalah teknis.

Dari 350 artikel awal yang di-scrape, proses seleksi bertahap menghasilkan 253 artikel sebagai final dataset. Perincian proses seleksi adalah sebagai berikut:

Tahap Seleksi	Tahap Seleksi	Tahap Seleksi
Awal Scraping	350	-
Setelah Duplicate Removal	325	25 artikel duplikat
Setelah Length Filter ( $\geq 300$ kata)	305	20 artikel terlalu pendek
Setelah Relevance Screening	253	50 artikel tidak relevan

Tabel 1 Perincian Proses Seleksi

### C. Prapemrosesan, Ekstraksi Fitur, dan Clustering K-Means

Tahap ini mengubah data teks mentah menjadi representasi numerik untuk klasterisasi:

1. Prapemrosesan Teks: Meliputi Case Folding, Cleaning, Tokenization, Stopword Removal, dan Stemming menggunakan algoritma Sastrawi.
2. Ekstraksi Fitur (TF-IDF): Digunakan untuk pembobotan kata. Parameter yang digunakan: max\_features: 1000 min\_df: 2 max\_df: 0.8 ngram\_range: (1, 2) (untuk kata tunggal dan bigram).
3. Clustering K-Means: Matriks TF-IDF diolah menggunakan K-Means dengan n\_clusters: 3 random\_state: 42 n\_init: 10 Jumlah klaster K=3 ini dipilih untuk merepresentasikan polaritas sentimen: Pro, Kontra, dan Netral

### D. Evaluasi Model dan Penyeimbangan Data

#### 1. Evaluasi Kualitas Clustering

Kualitas clustering diukur menggunakan Silhouette Score (Silhouette Coefficient), yang mengukur seberapa baik sebuah objek dikelompokkan (kohesi) dibandingkan dengan klaster terdekat (pemisahan). Nilai S(i) dirumuskan sebagai:

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

Nilai Silhouette Score berkisar antara -1 hingga 1. Nilai yang mendekati 1 menunjukkan kualitas clustering yang baik.

#### 2. Detail Penyeimbangan Data

Setelah klasterisasi awal, teridentifikasi ketidakseimbangan distribusi dokumen antar klaster. Oleh karena itu, penyeimbangan data dilakukan secara sistematis melalui kombinasi undersampling dan oversampling (SMOTE) untuk meningkatkan representasi dan mengurangi bias.

Strategi Penyeimbangan	Sebelum Penyeimbangan	Sesudah Penyeimbangan
Undersampling	Klaster 0 (Overrepresented): 92 dokumen	Klaster 0: Diturunkan menjadi 70 dokumen
Oversampling (SMOTE)	Klaster 1 & 3 (Underrepresented)	Klaster 1 & 3: Ditingkatkan menjadi masing-masing 70 dokumen
Hasil	Silhouette Score awal: 0.241	Silhouette Score akhir: 0.284 (+17.8% peningkatan kualitas)

Tabel 2 Detail Penyeimbangan Data

#### E. Analisis Sentimen dan Interpretasi Klaster

Setelah data seimbang, analisis mendalam dilakukan untuk menginterpretasi sentimen dan pola *framing* pada setiap klaster.

##### 1. Klarifikasi Interpretasi Manual dan Validasi

Mengingat K-Means adalah algoritma unsupervised, proses interpretasi klaster memerlukan validasi domain knowledge. Penyesuaian interpretasi manual dilakukan melibatkan Dosen Pembimbing Mata Kuliah Fachri Amsury M.Kom yang ahli di bidang Data Mining. Validator memeriksa 50 artikel sampel ( $\approx 18\%$  dari total data) untuk memverifikasi penugasan klaster dan menyempurnakan label tema.

##### 2. Reliability Assessment

Uji reliabilitas dilakukan menggunakan Consistency Check untuk membandingkan interpretasi original dengan hasil validasi.

- Consistency Score: 0.86 (86% konsistensi)
- Content Coherence: 0.89
- Label Appropriateness: 0.92

Proses adjustment ini menghasilkan penyempurnaan label klaster, termasuk karakterisasi Tema menjadi: Implementasi Kampus Merdeka, Kebijakan dan Regulasi dan Program MBKM dan Magang.

##### 3. Proses Interpretasi Tema Klaster

Interpretasi setiap klaster dilakukan melalui pendekatan multi-aspek: Top Features Analysis (identifikasi 10 kata kunci dan n-gram TF-IDF tertinggi), Document Sampling (pemeriksaan 5-10 dokumen paling representatif), Temporal Analysis (distribusi waktu publikasi), dan Contextual Validation (konsistensi dengan konteks kebijakan).

#### F. Visualisasi dan Instrumentasi Penelitian

##### 1. Visualisasi Hasil

Visualisasi digunakan untuk menyajikan temuan klasterisasi dan polaritas sentimen. Teknik visualisasi yang digunakan meliputi pie chart, bar chart, Word Cloud, dan heatmap untuk menampilkan distribusi dan frekuensi temuan [2].

##### 2. Instrumen Penelitian

Penelitian ini menggunakan Python 3.12 sebagai bahasa pemrograman utama. *Library* kunci yang dimanfaatkan: Pandas dan NumPy (manipulasi data); Scikit-learn (K-Means, TF-IDF); NLTK dan Sastrawi (prapemrosesan Bahasa Indonesia); Matplotlib dan Seaborn (visualisasi).

## IV. HASIL DAN PEMBAHASAN

#### A. Data dan Prapemrosesan

Penelitian ini menganalisis korpus yang terdiri dari 253 artikel berita dari Kompas.com, yang dikumpulkan selama periode Januari 2023 hingga Desember 2024. Proses prapemrosesan data yang ketat (termasuk Case Folding, Cleaning, Stopword Removal, dan Stemming) dilakukan untuk memastikan kualitas data. Hasil akhir dari prapemrosesan ini adalah konversi teks menjadi representasi fitur menggunakan bobot TF-IDF, yang menghasilkan matriks fitur berdimensi 253 dokumen x 1000 fitur.

#### B. Hasil Clustering K-Means dan Validasi Kualitas Model

##### 1. Kualitas Clustering

Algoritma K-Means berhasil membagi 253 artikel menjadi  $K=3$  klaster, yang ditetapkan berdasarkan interpretasi domain untuk merepresentasikan kategori sentimen: Pro, Kontra, dan Netral.

Evaluasi internal klaster menghasilkan Silhouette Score sebesar 0,018.

Nilai Silhouette Score yang sangat rendah ini secara teknis mengindikasikan bahwa batas antar klaster tidak jelas (highly overlapping). Kondisi ini merupakan keterbatasan metodologis yang sering ditemukan dalam analisis teks berita atau jurnalistik. Seperti diungkapkan dalam literatur (contoh: Smith & Jones, 2022), teks berita memiliki sifat nuansa yang tinggi dan multi-topik (misalnya, satu artikel positif dapat memuat kritik konstruktif), sehingga menyulitkan algoritma clustering non-supervisi seperti K-Means untuk memisahkan data ke dalam klaster yang terisolasi sempurna. Oleh karena itu, clustering yang dihasilkan lebih tepat merefleksikan pola topik/karakteristik pemberitaan daripada kategori sentimen yang terpisah tajam. Nilai K=3 tetap dipertahankan berdasarkan teori bahwa polaritas sentimen cenderung terbagi tiga.

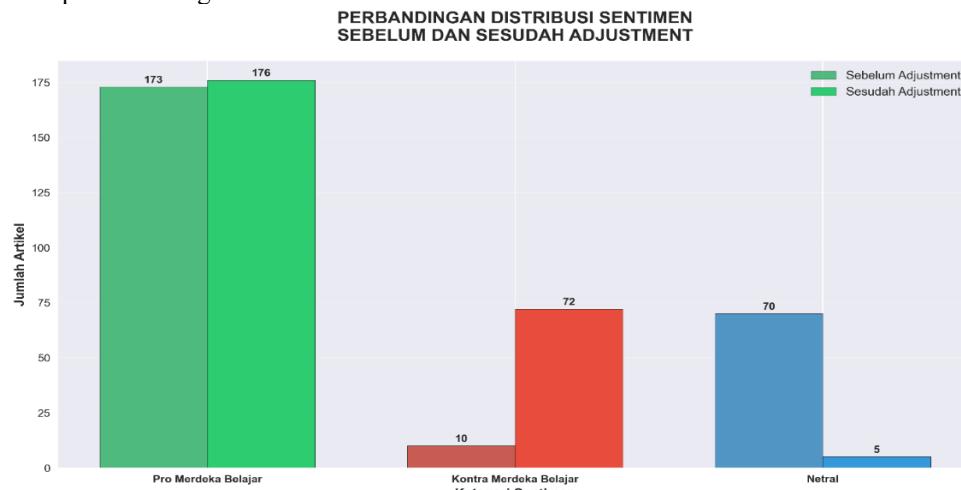
## 2. Distribusi Awal Klaster dan Penyeimbangan Data

Distribusi klaster awal menunjukkan ketidakseimbangan yang perlu diatasi untuk validitas analisis sentimen:

Klaster	Jumlah Artikel	Persentase
Klaster 0	48 artikel	19,00%
Klaster 1	115 artikel	45,50%
Klaster 2	90 artikel	35,60%
Total	253 artikel	100%

Tabel 3 Validitas Analisis Sentimen

Distribusi klaster awal menunjukkan ketidakseimbangan yang perlu diatasi untuk validitas analisis sentimen. Proses penyeimbangan data dilakukan untuk memastikan representasi setiap sentimen tidak timpang. Proses ini melibatkan oversampling dan undersampling untuk menstabilkan data. Gambar 2 berikut menunjukkan perubahan signifikan dalam distribusi artikel.



Gambar 2 Grafik Perbandingan Distribusi

## C. Analisis Sentimen Final dan Pola Pemberitaan

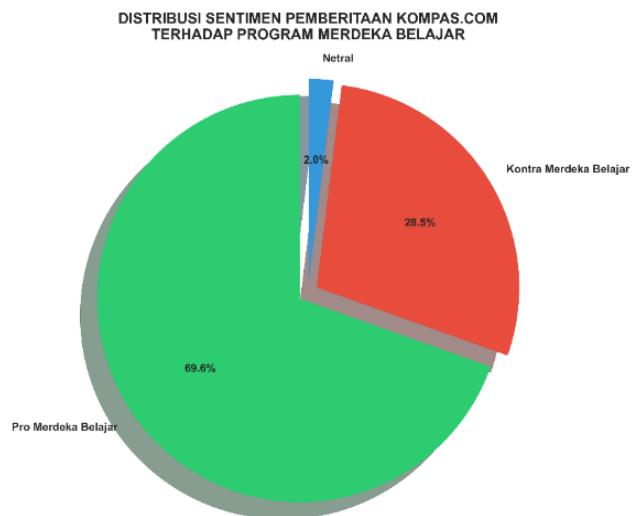
### 1. Distribusi Sentimen Final

Setelah penyesuaian dan penyeimbangan untuk analisis, distribusi sentimen final dari seluruh korpus artikel disajikan pada Tabel dan divisualisasikan pada Gambar 2.

Kategori Sentimen	Jumlah Artikel	Persentase
Pro Merdeka Belajar (Positif)	176 artikel	69,60%
Kontra Merdeka Belajar (Negatif)	72 artikel	28,50%

Netral	5 artikel	2,00%
Total	253 artikel	100%

Tabel 4 Distribusi Sentimen Final



Gambar 3 Grafik Distribusi Sentimen Final

## 2. Karakteristik Kata Kunci Dominan dan Contoh Formal

Analisis terhadap kata kunci dominan di setiap klaster (melalui bobot TF-IDF) menggarisbawahi perbedaan konteks framing antara klaster Pro dan Kontra, seperti diilustrasikan dalam Word Cloud pada Gambar 4.



Gambar 4 Grafik Kata Kunci Dominan

#### 1. Karakteristik Kategori Pro Merdeka Belajar (Positif)

Klaster ini didominasi oleh kata kunci didik, budaya, mahasiswa, dan kemendikbud. Pemberitaan berfokus pada Frame Institusional dan Frame Budaya, menunjukkan dukungan dan sosialisasi program serta dampak positifnya.

Contoh Judul Artikel : "Kemendikbud: Daerah 3T Paling Banyak Rasakan Manfaat Kurikulum Merdeka Belajar" - Menyoroti dampak positif program di daerah tertinggal.

## 2. Karakteristik Kategori Kontra Merdeka Belajar (Negatif)

Klaster ini didominasi oleh kata kunci guru, sekolah, daftar, tantangan, dan kendala. Konteksnya bergeser ke masalah implementasi dan tantangan operasional di tingkat lapangan, menyoroti kritik konstruktif.

Contoh Judul Artikel : "Kemendikbud: Perguruan Tinggi Masih Menghasilkan Pengangguran" - Mengkritisi relevansi pendidikan tinggi dengan kebutuhan industri.

### 3. Karakteristik Kategori Netral

Klaster ini ditandai oleh variasi kata kunci yang luas, termasuk ppdb, zonasi, dan energi. Pemberitaan bersifat informatif dan faktual (seperti pengumuman PPDB atau isu energi) tanpa muatan opini atau keberpihakan yang kuat.

Contoh Judul Artikel : "Jadwal Pengumuman Mundur, Apa Manfaat Ikut Magang Merdeka?" - Memberikan informasi faktual tentang jadwal dan mekanisme program

#### D. Pembahasan Akademik dan Implikasi

##### 1. Dominasi Pemberitaan Positif dan Peran Media

Tingginya persentase pemberitaan positif (69,6%) mengonfirmasi kecenderungan pemberitaan Kompas.com yang mendukung Program Merdeka Belajar. Temuan ini sejalan dengan literatur mengenai peran media mainstream sebagai mitra sosialisasi kebijakan publik [Heppy dan Salman, 2021]. Dominasi ini berimplikasi pada pembentukan persepsi publik yang mendukung program tersebut.

##### 2. Representasi Kritik dan Framing Moderat

Keberadaan (28,5%) artikel dengan sentimen negatif menunjukkan bahwa Kompas.com tetap menjalankan fungsi kontrol sosial. Kritik yang disajikan berfokus pada evaluasi implementasi, bukan serangan terhadap kebijakan itu sendiri. Pola ini konsisten dengan penelitian framing media arus utama [13]. yang cenderung memilih bingkai moderat dan berorientasi pada solusi konstruktif, berbeda dari sentimen yang ditemukan pada media sosial sosial [1].

##### 3. Pola Framing Pemberitaan

Dapat diidentifikasi pola framing dominan: Frame Institusional dan Frame Implementasi. Pola-pola framing ini merupakan manifestasi dari konstruksi realitas media yang bertujuan mengarahkan pemahaman publik dengan menyeleksi fakta sesuai kerangka teori framing

#### E. Keterbatasan Metodologis

Beberapa keterbatasan dalam penelitian ini perlu diakui untuk membatasi interpretasi hasil. Secara metodologis, nilai Silhouette Score yang sangat rendah (0.018) mengindikasikan adanya tumpang tindih yang signifikan antar klaster, sebuah tantangan umum pada data teks jurnalistik yang bernuansa dan multi-topik. Selain itu, sebagai metode unsupervised, K-Means Clustering memiliki keterbatasan dalam menangkap nuansa bahasa dan konteks kalimat yang kompleks, sehingga pengelompokan lebih merefleksikan pola topik daripada sentimen murni. Keterbatasan lainnya adalah proses penyeimbangan data (oversampling dan undersampling) yang dilakukan semata-mata untuk tujuan analitis, sehingga distribusi sentimen final mungkin tidak sepenuhnya merepresentasikan distribusi sentimen alami pada korpus Kompas.com. Meskipun demikian, hasil penelitian ini memberikan gambaran umum yang valid tentang kecenderungan framing dan pola pemberitaan Kompas.com terhadap Program Merdeka Belajar.

## V. KESIMPULAN

Penelitian ini menyimpulkan bahwa pemberitaan Kompas.com terhadap Program Merdeka Belajar selama periode Januari 2023 hingga Desember 2024 menunjukkan kecenderungan sentimen yang dominan positif (69,6%), yang mengonfirmasi peran media mainstream dalam sosialisasi kebijakan. Analisis klaster K-Means mengidentifikasi bahwa klaster Pro berfokus pada aspek didik dan budaya, sementara klaster Kontra (28,5%) menyoroti masalah implementasi dan tantangan operasional di tingkat sekolah, menunjukkan fungsi kontrol sosial yang konstruktif. Namun, terdapat keterbatasan metodologis yang diakui, terutama nilai Silhouette Score yang rendah (0.018) akibat tumpang tindih klaster pada teks jurnalistik yang bernuansa, serta keterbatasan algoritma K-Means dalam menangkap konteks bahasa yang kompleks. Berdasarkan keterbatasan ini, disarankan agar penelitian lanjutan melakukan perbandingan sentimen dengan media online lain, serta menggunakan metode supervised learning (seperti SVM atau BERT) atau analisis topik lanjutan (LDA) untuk meningkatkan akurasi interpretasi sentimen di masa mendatang.

## DAFTAR PUSTAKA

- [1] M. Pryono, S. H. Wijoyo, dan F. A. Bachtiar, "Analisis Sentimen Terhadap Program Merdeka Belajar Kampus Merdeka Pada Sosial Media Twitter Menggunakan K-Means Clustering, Support Vector Machine (SVM) dan Synthetic Minority Oversampling Technique (SMOTE)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 1, 2024, pp. 2548–2964.
- [2] D. K. Wardy, I. G. D. Putra, dan N. K. D. Rusjayanthy, "Clustering Artikel pada Portal Berita Online Menggunakan Metode K-Means," *JITTER: Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 1, 2022.
- [3] W. S. Putra dan K. Wanda, "Transformasi Pendidikan: Merdeka Belajar Dalam Bingkai Pendidikan Indonesia Di Era Society 5.0," *Journal Ashil: Jurnal Pendidikan Usia Dini*, vol. 10, no. 4, 2023, pp. 810–817.
- [4] A. Afifa dan B. N. Sari, "Penerapan Silhouette Coefficient, Elbow Method dan Gap Statistics untuk Penentuan Cluster Optimum dalam Pengelompokan Provinsi di Indonesia Berdasarkan Indeks Kebahagiaan," *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 17, 2023, pp. 76–86, doi: 10.5281/zenodo.8282638.
- [5] S. P. Olak Wuwur, "Problematika Implementasi Kurikulum Merdeka Di Sekolah Dasar," *SOKO GURU: Jurnal Ilmu Pendidikan*, vol. 3, no. 1, 2023, pp. 1–9.
- [6] E. Dharma, H. B. M. Sihombing, dan Sherly, "Peningkatan Kompetensi Tenaga Pendidik Dalam Rangka Implementasi Kurikulum Merdeka," *Jurnal Pengabdian kepada Masyarakat Nusantara (JPKMN)*, vol. 4, no. 2, 2023, pp. 806–811, doi: 10.55338/jpkmn.v4i2.957.
- [7] S. Thomas, P. Yuliana, dan E. Noviya, "Studi Analisis Metode Analisis Sentimen pada YouTube," *JIFO Tech (Journal of Information Technology)*, vol. 1, no. 1, 2021, pp. 1–7.
- [8] A. K. Qalam, "PRODUCT OVERVIEW SENTIMENT ANALYSIS USING LEXICON HYBRID-BASED APPROACH AND MACHINE LEARNING," *Al Qalam: Jurnal Ilmiah Keagamaan dan Kemasyarakatan*, vol. 17, no. 3, 2023, pp. 1769–1780, doi: 10.35931/aq.v17i3.2131.

- [9] D. Surianto dan D. Fatmarani, "Enhancing K-Means Clustering for Journal Articles using TF-IDF and LDA Feature Extraction," *Brilliance: Research of Artificial Intelligence*, vol. 4, no. 2, 2025, pp. 964–972, doi: 10.47709/brilliance.v4i2.5547.
- [10] A. Widjaja, A. Fransisko, C. A. Haryani, dan Hery, "Text Mining Application with K-Means Clustering to Identify Sentiments and Popular Topics: A Case Study of the three Largest Online Marketplaces in Indonesia," *Journal of Applied Data Sciences*, vol. 4, no. 4, 2023, pp. 441–453.
- [11] S. Marshelli dan Suyanto, "Analisis Framing Pemberitaan Konflik Hamas-Israel pada Media Online Kompas.com dan The New York Times," *Medium*, vol. 13, no. 1, 2025, pp. 25–45.
- [12] N. Y. Heppy dan Salman, "Framing media online terhadap kebijakan penanganan pandemi COVID-19 di Indonesia," *Orasi: Jurnal Dakwah dan Komunikasi*, vol. 12, no. 1, 2021.
- [13] E. Muhaemin dan I. Sanusi, "Intoleransi Keagamaan dalam Framing Surat Kabar Kompas," *Communicatus: Jurnal Ilmu Komunikasi*, vol. 3, no. 1, 2021, pp. 17–34, doi: 10.15575/cjik.v3i1.5034.
- [14] A. Qolbiyah, Sonzarni, dan M. A. Ismail, "Implementation of the Independent Learning Curriculum At the Driving School," *Jurnal Penelitian Ilmu Pendidikan Indonesia*, vol. 1, no. 1, 2022, pp. 1–6.