

KLASTERISASI TEKS BERITA PEMILU UNTUK ANALISIS FRAMING DI KOMPAS.COM

Atik Ariesta^{*1}, Jeanie Annissa²

^{1,2}Universitas Budi Luhur, Jakarta Selatan

Email: ¹atik.ariesta@budiluhur.ac.id, ²jeanie.annissa@budiluhur.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 10 Desember 2024, diterima untuk diterbitkan: 30 Oktober 2025)

Abstrak

Penelitian ini bertujuan untuk menganalisis *framing* berita pemilu di Kompas.com dengan mengelompokkan berita berdasarkan pola teks menggunakan algoritma *K-Means*. Kompas.com sebagai salah satu media *online* memiliki peran dalam membentuk opini publik pada masa pemilu, namun pendekatan kuantitatif untuk memetakan *framing* media *online* masih jarang dilakukan. Penelitian ini akan menerapkan metode klusterisasi untuk menganalisis *framing* berita. Data dikumpulkan melalui *web scrapping* dari Juni 2023 hingga Januari 2024, menghasilkan 6.570 berita. Hasil pembobotan kata menggunakan TF-IDF digunakan untuk klusterisasi, dengan jumlah kluster optimal ditentukan oleh *elbow method*. Hasil klusterisasi menunjukkan *framing* pemberitaan setiap bulan, yang di evaluasi menggunakan Davies-Boudin Indeks (DBI) dengan menunjukkan hasil kluster yang layak. Analisis *framing* dilakukan pada kluster dengan data terbanyak tiap bulannya, menggunakan dua dimensi (seleksi isu dan penonjolan aspek), serta empat elemen (*framing: define problem, diagnose cause, make moral judgment, dan treatment recommendation*). Hasil penelitian menunjukkan bahwa pola *framing* yang muncul mencerminkan dinamika pemilu, seperti isu koalisi, dukungan calon presiden, serta kontroversi regulasi pemilu. Temuan ini memberikan kontribusi tentang bagaimana media menyajikan berita politik serta memperkuat pentingnya klusterisasi dalam studi *framing* media.

Kata kunci: klusterisasi; teks; *framing*; berita; pemilu

TEXT CLUSTERING OF ELECTION NEWS FOR FRAMING ANALYSIS IN KOMPAS.COM

Abstract

This study aims to analyze the framing of election news on Kompas.com by grouping news based on text patterns using the K-Means algorithm. Kompas.com as one of the online media has a role in shaping public opinion during the election period, but a quantitative approach to mapping online media framing is still rare. This study will apply the clustering method to analyze news framing. Data were collected through web scrapping from June 2023 to January 2024, resulting in 6,570 news items. The results of word weighting using TF-IDF were used for clustering, with the optimal number of clusters determined by the elbow method. The clustering results show news framing each month, which is evaluated using the Davies-Boudin Index (DBI), showing decent cluster results. Framing analysis was carried out on the cluster with the most monthly data, using two dimensions (issue selection and aspect highlighting), and four elements (*framing: define problem, diagnose cause, make moral judgment, and treatment recommendation*). The study results show that the emerging framing patterns reflect the election dynamics, such as coalition issues, presidential candidate support, and election regulation controversies. These findings contribute to how the media presents political news and strengthen the importance of clustering in media framing studies.

Keywords: clustering; text; *framing*; news; election

1. PENDAHULUAN

Pemilihan umum (Pemilu) adalah salah satu momen penting dalam kehidupan politik suatu negara. Media berita memiliki peran krusial dalam membentuk persepsi masyarakat terkait dengan kandidat, isu-isu politik, dan hasil pemilu. *Framing* media berita, yaitu cara media menyajikan informasi, dapat memengaruhi cara pemilih memahami dan mengevaluasi proses pemilu dan pemimpin yang terpilih. Kompas.com sebagai salah satu situs berita terkemuka di Indonesia, memiliki pengaruh besar dalam membentuk opini publik. Kompas.com menyediakan platform khusus untuk pemberitaan seputar pemilu yaitu pemilu.kompas.com. Berdasarkan Similarweb jumlah kunjungan (*traffic*) website Kompas.com berada pada peringkat ketiga kategori News & Media Publisher (Similarweb, 2023).

Dalam era digital, berita pemilu yang disajikan oleh Kompas.com dan media serupa tidak hanya tersedia dalam bentuk cetak, tetapi juga dalam bentuk teks digital. Analisis teks berita pemilu dapat menjadi alat untuk memahami *framing* media. Klusterisasi teks adalah metode yang dapat digunakan untuk mengidentifikasi pola dalam data teks, yang dapat membantu untuk memahami cara media berita memproses informasi pemilu. Salah satu algoritma klusterisasi yang populer adalah K-Means.

K-Means adalah teknik klusterisasi yang telah digunakan dalam berbagai bidang, termasuk analisis teks. Algoritma K-Means merupakan algoritma Klusterisasi tanpa memerlukan supervisi (*unsupervised*) (Afida dkk., 2021). Pada penelitian ini, K-Means akan digunakan untuk mengelompokkan artikel berita pemilu berdasarkan pola-pola yang ada dalam teks. Hasil klusterisasi ini dapat membantu mengidentifikasi pola *framing* media yang digunakan dalam melaporkan berita pemilu. Analisis *framing* dilakukan menjadi dua tahap yaitu dimensi *framing* dan elemen *framing*. Dimensi *framing* terdiri dari seleksi isu dan penonjolan aspek. Elemen *framing* akan menganalisis berdasarkan pemilihan kata, kalimat, dan gambar yang digunakan. Elemen *framing* terdiri dari 4 empat model yaitu *define problem*, *diagnose cause*, *moral judgment*, dan *treatment recommendation* (Hidayah & Riauan, 2021).

Penelitian (Purniawan dkk., 2022) membahas *clustering* berita dengan algoritma TF-IDF dan K-Means, menggunakan data *crawling* dari detik.com. Tujuannya adalah mengidentifikasi kelompok berita melalui implementasi K-Means berdasarkan pembobotan kata dengan TF-IDF. Dengan 124.509 berita dari 2018, hasilnya menunjukkan 27 kelompok berita yang berhasil diidentifikasi, menunjukkan kemampuan aplikasi dalam memproses data besar. Penelitian (Rheza & Metandi, 2020) menggunakan K-Means Clustering untuk mengelompokkan jenis komentar pada *tweet*

PSSI, membedakan ironi, sinisme, dan sarkasme. Dari klusterisasi, ditemukan 20 data sinisme di *cluster* 1, 3 data sarkasme di *cluster* 2, dan 27 data ironi di *cluster* 3. Penelitian (Simanjuntak dkk., 2023) membahas analisis kluster pada data teks dari media sosial, khususnya Twitter, menggunakan K-Means Clustering. Hasilnya menunjukkan kluster dominan terkait ekonomi, politik, sosial, dan olahraga. Analisis dilakukan dengan Orange Data Mining, termasuk analisis sentimen dengan *lexicon* terintegrasi.

Penelitian (Agustiani dkk., 2020) menganalisis bingkai berita Pemilihan Presiden Indonesia 2019 di Zonasultra.com, menyoroti preferensi terhadap pasangan calon 01. Penelitian (Hasan & Fathan, 2020) pada Solopos.com mengungkap *framing* politik media, cenderung memihak PDIP menjelang Pilkada 2020. Sementara Penelitian (Siregar & Qurniawati, 2022) fokus pada *buzzers* dalam Pemilu 2019 di tempo.co, menunjukkan perhatian terhadap sisi negatif *buzzer* sebagai sumber propaganda politik yang perlu diatur. Tiga penelitian ini memberikan wawasan mendalam terkait *framing* dan politik media dalam konteks pesta politik di Indonesia.

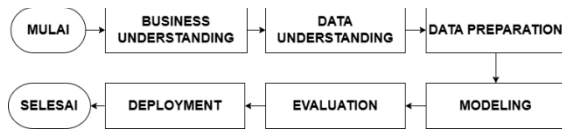
Penelitian sebelumnya yang dilakukan oleh Purniawan dkk. (2022) dan Simanjuntak dkk. (2023) telah memanfaatkan k-means untuk melakukan klusterisasi teks pada berita daring maupun media sosial, namun klusterisasi digunakan untuk identifikasi topik dan tema dominan tanpa mengaitkan bagaimana media menyajikan informasi. Penelitian yang membahas tentang *framing* media oleh Agustiani dkk. (2020) dan Hasan & Fathan (2020) memberikan analisis kualitatif terhadap *framing* media dalam memberitakan kandidat politik, tetapi belum menggunakan klusterisasi dalam proses pengelompokannya.

Penelitian ini bertujuan untuk melakukan klusterisasi berita pemilu pada media Kompas.com menggunakan algoritma K-Means dan mengaitkannya dengan analisis dimensi dan elemen *framing* untuk melihat pembentukan berita. Berbeda dengan penelitian sebelumnya yang melakukan analisis topik dan *framing* secara terpisah, penelitian ini menggabungkan keduanya dalam penelitian yang terstruktur. Pemilihan Kompas.com sebagai objek penelitian dapat memberikan pemahaman peran media online dalam pembentukan opini publik. K-Means dipilih karena dapat menangani data besar secara efisien dan dapat menghasilkan kluster untuk dianalisis secara semantik. Penggabungan dua hal ini K-Means dan *framing* diharapkan dapat menjadi penelitian kualitatif dalam penelitian *framing* berbasis teks.

2. METODE PENELITIAN

Gambar 1 merupakan langkah-langkah penelitian mulai dari *Business Understanding*, *Data*

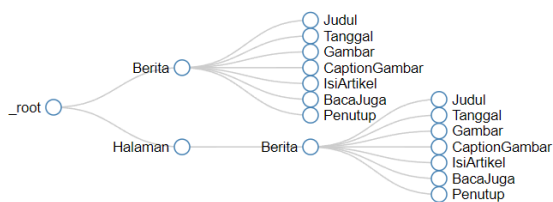
Understanding, Data Preparation, Modeling, Evaluation, dan Deployment.



Gambar 1. Langkah Penelitian CRISP-DM

Langkah pertama penelitian ini adalah *business understanding*. Pada langkah ini peneliti menentukan tujuan dari penelitian dan identifikasi permasalahan yang dipecahkan. Pada penelitian ini identifikasi permasalahan adalah Klasterisasi berita pemilu untuk analisis *framing* berita pemilu pada website pemilu.kompas.com. *Framing* dilakukan untuk mengetahui bagaimana media massa secara aktif memilih dan menonjolkan aspek-aspek tertentu sehingga dapat membentuk persepsi publik (Yazmi dkk., 2024). Pada langkah ini akan menghasilkan penetapan masalah utama penelitian terkait klasterisasi teks berita pemilu untuk analisis *framing*.

Langkah kedua adalah *data understanding*. Pengumpulan berita pemilu dari pemilu.kompas.com menggunakan web *scraping*. Web *Scraping* adalah teknik pengambilan data dan informasi pada suatu halaman web (Ayub, 2021) (Sahria Yoga, 2020). Tools yang digunakan adalah web *scraper* yang merupakan *add-ons* pada Google Chrome. Data yang akan dikumpulkan adalah berita antara Juni 2023 hingga Januari 2024. Pemilihan rentang waktu ini didasarkan pada pertimbangan bahwa periode tersebut merupakan masa menjelang pemilihan presiden dan wakil presiden, yang ditandai dengan eskalasi isu politik, strategi kampanye, dan tingginya perhatian media terhadap proses demokrasi.



Gambar 2. Selector Graph Web Scraper

Gambar 2 merupakan *selector graph* konfigurasi pada web *scraper*. Setiap node berisi selector/tag HTML dari berita yang akan *discraping*. *_root* merupakan titik awal melakukan *scraping* yang terdiri dari dua bagian yaitu Berita dan Halaman. Berita merupakan *selector* untuk *scraping* list berita yang tampil pada halaman pertama di pemilu.kompas.com, Halaman merupakan *selector* untuk *scraping* list berita yang tampil pada halaman kedua dan seterusnya di pemilu.kompas.com. Judul, Tanggal, Gambar, *CaptionGambar*, Isi Artikel, BacaJuga, Penutup merupakan *selector* detail berita yang akan *discraping*. Luaran dari langkah ini adalah Dataset berita pemilu yang terkumpul dari

pemilu.kompas.com antara Juni 2023 hingga Januari 2024.

Langkah ketiga adalah *data preparation*. Peneliti melakukan pemrosesan teks yang terdiri dari *tokenization*, *stopword removal*, dan *stemming* (Jo, 2019). Teks yang akan dilakukan pemrosesan adalah Judul Berita. Kegiatan yang dilakukan pada saat *tokenization* adalah menghapus spesial karakter dan angka, mengubah ke huruf kecil, dan memecah teks berdasarkan *whitespace* (spasi). Pada saat *stopword removal* adalah menghilangkan kata-kata yang tidak relevan yang sering muncul pada teks, preposisi, dan konjungsi. *Stemming* dilakukan berdasarkan hasil dari *stopword removal*. *Stemming* akan mengubah setiap *token* yang memiliki imbuhan menjadi bentuk akar kata. *Stemming* menggunakan Sastrawi yang merupakan *stemmer* paling sesuai untuk Bahasa Indonesia (Firman dkk., 2022). Selanjutnya setelah pemrosesan data, peneliti akan melakukan pembobotan kata dengan menggunakan Term *Frequency-Invers Document Frequency* (TF-IDF). Luaran dari langkah ketiga adalah Judul berita yang telah melalui tahapan pemrosesan teks serta matriks bobot TF-IDF untuk setiap kata dalam judul berita.

Langkah keempat adalah *modeling*. Peneliti melakukan klasterisasi dari judul teks yang sudah mendapatkan nilai bobotnya. Sebelum melakukan proses Klasterisasi ditentukan nilai *k* yang optimal dengan menggunakan *elbow method*. Bentuk siku dari grafik *elbow method* merupakan nilai *k* yang akan digunakan untuk menentukan nilai klaster. Luaran dari langkah ini adalah Jumlah klaster yang optimal ditentukan menggunakan *elbow method*, dan klasterisasi judul berita dilakukan berdasarkan nilai TF-IDF. Algoritma K-Means yang bekerja dengan cara menginisialisasi *k* pusat klaster secara acak dan kemudian mengalokasikan setiap data ke klaster berdasarkan kedekatannya dengan pusat klaster terdekat (Widaningrum dkk., 2022).

Langkah kelima adalah evaluasi. Evaluasi hasil klasterisasi menggunakan *Davies-Bouldin Index* (DBI) sebagai langkah untuk mengukur kualitas dan kecocokan klaster yang dihasilkan oleh algoritma K-Means. DBI untuk suatu klaster tertentu dihitung sebagai rasio dari *Intra-Cluster Dispersion* dengan Jarak Antar Klaster dengan menggunakan *Euclidean Distance* (Anita & Aliando, 2022). Hasil dari langkah ini adalah nilai DBI dari setiap dataset.

Langkah keenam adalah *deployment*. Pada langkah ini peneliti akan memilih klaster dengan jumlah data terbanyak disetiap bulannya. Dari klaster yang terpilih peneliti menentukan dimensi *framing* dan elemen *framing*. Dimensi *framing* yaitu pemilihan isu dan penonjolan aspek. Elemen *framing* yaitu *define problem*, *diagnose cause*, *make moral judgment*, dan *treatment recommendation*. *Define problem* merupakan elemen untuk mendefinisikan masalah terkait bagaimana suatu isu penelitian dapat dilihat, sebagai masalah apa. *Diagnose cause* merupakan elemen untuk

memperkirakan sumber masalah terkait peristiwa yang dilihat disebabkan oleh apa dan mengaskan aktor yang dianggap sebagai penyebab masalah. *Make moral judgment* merupakan elemen untuk membuat keputusan moral untuk menjelaskan masalah atau nilai apa yang digunakan untuk mendelegitimasi suatu tindakan. *Treatment recommendation* merupakan elemen untuk menyelesaikan masalah, penyelesaian apa yang ditawarkan untuk mengatasi masalah isu dan atau jalan apa yang ditawarkan untuk mengatasi masalah (Boer dkk., 2020). Luaran dari langkah ini adalah dimensi *framing* dan elemen *framing* pada berita pemilu.

3. HASIL DAN PEMBAHASAN

4. 3.1. Pengumpulan Data

Tabel 1 merupakan jumlah berita yang diperoleh dari media *online* Kompas. Secara umum, jumlah berita memiliki tren peningkatan tiap bulannya, walaupun terjadi penurunan pada bulan Oktober 2023 (807 berita) dibandingkan dengan bulan September (812 berita). Total keseluruhan berita hasil *web scrapping* adalah 6.570.

No	Bulan / Tahun	Jumlah Berita
1	Juni 2023	571
2	Juli 2023	750
3	Agustus 2023	767
4	September 2023	812
5	Oktober 2023	807
6	November 2023	863
7	Desember 2023	977
8	Januari 2024	1.023
Total		6.570

merupakan Atribut yang dihasilkan dari *web scrapping* pada web Kompas.com yaitu pemilu.kompas.com. Terdapat 8 atribut yang diperoleh menggunakan *tools web scrapper google chrome*.

Tabel 3 adalah atribut yang ditentukan untuk digunakan dalam proses Klasterisasi yaitu Tanggal Berita, Judul Berita, Link Berita, Link Gambar Berita, dan Isi Berita. Atribut dipilih sesuai dengan kebutuhan untuk analisis *framing* berita. Klasterisasi menggunakan atribut judul berita karena judul sering memuat kata-kata kunci dan sudut pandang media sebagai indikator arah *framing*. Oleh karena itu, judul sangat relevan untuk membantu proses identifikasi topik dan elemen *framing* dalam analisis.

Pada penelitian ini akan dihasilkan 8 *dataset* berdasarkan data berita di setiap bulan (

Tabel 1). Atribut yang digunakan sesuai yang disajikan pada Tabel 3.

Tabel 2. Atribut *Scrapping*

No	Atribut	Keterangan
1	<i>web-scrapper-order</i>	Urutan <i>tools web scrapper</i> melakukan <i>scrapping</i> data berita
2	<i>web-scrapper-start-url</i>	Link url awal <i>web scrapper</i> melakukan <i>scrapping</i> data berita
3	Berita	Berisi tanggal dan judul berita dari berita yang <i>discrapping</i>
4	berita-href	Link url dari artikel/berita yang <i>discrapping</i>
5	Judul	Judul berita dari artikel yang <i>discrapping</i>
6	Tanggal	Tanggal berita dari artikel yang <i>discrapping</i>
7	GambarBerita-src	Link gambar utama dari berita yang <i>discrapping</i>
8	CaptionGambar	Teks singkat dari gambar utama yang <i>discrapping</i>
9	IsiArtikel	Isi berita dari berita yang <i>discrapping</i>
10	Halaman	Halaman dari berita yang <i>discrapping</i> pada list berita
11	Halaman-href	Link url dari berita yang <i>discrapping</i>

Tabel 3. Atribut yang Digunakan

No.	Atribut Dataset	Keterangan
1	Tanggal	Tanggal Berita
2	Judul	Judul Berita
3	Link Berita	Link Berita
4	Link Gambar	Link Gambar Utama Berita
5	Isi Berita	Isi Berita

Tabel 4 merupakan dataset berita Kompas bulan Juni 2023 sebagai contoh dari 8 dataset yang dihasilkan

3.2. Pemrosesan Teks

Pemrosesan Teks pada penelitian ini berawal dari *tokenization*, *stopword*, dan *stemming*. *Attribute* yang diproses untuk digunakan dalam penelitian adalah Judul. Tabel 6 merupakan contoh proses menghilangkan karakter spesial dan angka, mengubah menjadi huruf kecil (*lowercase*), dan *tokenization* (Jo, 2019).

Tabel 5 adalah pemrosesan teks dari judul awal yang kemudian akan di *transform* dengan menghilangkan karakter spesial dan angka, dan mengubahnya menjadi huruf kecil (*lowercase*). Disajikan pada kolom *Transform Teks* sudah tidak ada lagi karakter spesial seperti kutip dua (""), dan titik dua (:), serta sudah diubah menjadi huruf kecil semua. Pada kolom *Tokenization* merupakan hasil *tokenizing* yang memisahkan judul berita berdasarkan *whitespace* (spasi).

Tabel 6 hasil dari *Tokenization* akan dilakukan proses *stopword removal* seperti terlihat pada kolom *Stopword Removal* sudah tidak ada lagi kata 'untuk', 'cawecawe', dan 'dan' yang merupakan sebuah konjungsi. Pada *stemming* dilakukan juga penyatuan token menjadi satu kalimat.

Tabel 4. Dataset Penelitian Berita Kompas Bulan Juni 2023

Tanggal	Judul	Link Berita	Link Gambar	Isi Berita
01/06/2023	Kritik Jokowi, Puskapol UI: Dalih Cawecaw Pilpres untuk "Bangsa dan Negara" Alasan Klise	http://nasional.kompas.com/read/2023/06/01/17004851/kritik-jokowi-puskapol-ui-dalih-cawecaw-pilpres-untuk-bangsa-dan-negara	https://asset.kompas.com/crops/6hFfZVTDq71LRjLVQpLokBefDQ=/0x0:0x0/750x500/data/photo/2023/05/30/64754db5ac980.jpg	JAKARTA, KOMPAS.com - Pusat Kajian Politik Universitas Indonesia (Puskapol UI) menyinggung adanya problem serius dari kualitas politikus dalam negeri, merujuk pada pernyataan Presiden RI Joko Widodo yang mengaku tak akan netral pada Pilpres 2024 dan akan cawecawecaw. Sikap campur tangan kepala negara dalam urusan elektoral di mana ia tak menjadi salah satu kandidatnya dinilai menunjukkan rendahnya etika dan sikap kenegarawanan.....
30/06/2023	UU Pemilu Bolehkan Presiden hingga Wakil Bupati Ikut Kampanye, Simak Aturannya	http://nasional.kompas.com/read/2023/06/30/06570071/uu-pemilu-bolehkan-presiden-hingga-wakil-bupati-ikut-kampanye-simak	https://asset.kompas.com/crops/RQqmKfDOSv-1S4Uc6iljoZ7lbA=/0x0:1920x1280/750x500/data/photo/2022/10/28/635b51e75c2ae.jpg	JAKARTA, KOMPAS.com - Presiden hingga wakil bupati rupanya dapat melakukan kampanye atau bergabung sebagai tim kampanye. Hal ini termaktub dalam Undang-Undang (UU) Nomor 7 Tahun 2017 tentang Pemilihan Umum (Pemilu). Pada Pasal 281, UU Pemilu mengatur bahwa presiden, wakil presiden, menteri, gubernur, wakil gubernur, bupati, wakil bupati, Walikota, dan wakil Walikota boleh terlibat dalam kampanye peserta pemilu dengan sejumlah syarat. Syarat-syarat dalam Pasal 281 itu yakni:

Tabel 5. Transform dan Tokenization Judul Berita

Judul Awal	Transform Teks	Tokenization
Kritik Jokowi, Puskapol UI: Dalih Cawecaw Pilpres untuk "Bangsa dan Negara" Alasan Klise	kritik jokowi puskapol ui dalih cawecaw pilpres untuk bangsa dan negara alasan klise	['kritik', 'jokowi', 'puskapol', 'ui', 'dalih', 'cawecaw', 'pilpres', 'untuk', 'bangsa', 'dan', 'negara', 'alasan', 'klise']

Tabel 6. Stopword Removal Judul Berita

Tokenization	Stopword Removal
['kritik', 'jokowi', 'puskapol', 'ui', 'dalih', 'cawecaw', 'pilpres', 'untuk', 'bangsa', 'dan', 'negara', 'alasan', 'klise']	['kritik', 'jokowi', 'puskapol', 'ui', 'dalih', 'pilpres', 'bangsa', 'negara', 'alasan', 'klise']

Tabel 7. Stemming Judul Berita

Stopword Removal	Stemming
['kritik', 'jokowi', 'puskapol', 'ui', 'dalih', 'pilpres', 'bangsa', 'negara', 'alasan', 'klise']	kritik jokowi puskapol ui dalih pilpres bangsa negara alas klise

Tabel 7 hasil dari *stemming* dari *stopword removal* dan menyatukan setiap token menjadi kalimat. Setelah dilakukan *stemming* kata 'alasan' berganti menjadi kata dasarnya yaitu 'alas' seperti yang terlihat pada kolom *stemming*.

Langkah selanjutnya pada pemrosesan teks adalah melakukan pembobotan kata dari judul berita. Tabel 8 merupakan contoh hasil TF-IDF dari 3 berita yaitu Berita 1, Berita 2, dan Berita 3. Sel yang memiliki angka artinya Berita tersebut mengandung kata sesuai dengan kolom, sedangkan sel berisi 0 artinya Berita tersebut tidak memiliki kata sesuai dengan kolom. Berita 1 memiliki kata "ada", "anies", "demokrat", "kejut", "pacitan", "sby". Berita 2 memiliki kata "elektabilitas", "ganjar", "litbang", "menang", "prabowo", "salip", "santai", "sebentar", "survei". Berita 3 memiliki kata

"airlangga", "buka", "caleg", "ingin", "klaim", "pdip", "pemilu", "proporsional", "sistem".

Tabel 8. Term Frequency – Inverse Document Frequency (TF-IDF) Judul Berita

	Berita 1	Berita 2	Berita 3
ada	0,4041	0	0
airlangga	0	0	0,3842
anies	0,2317	0	0
buka	0	0	0,3394
caleg	0	0	0,3438
demokrat	0,2427	0	0
elektabilitas	0	0,2733	0
ganjar	0	0,1448	0
ingin	0	0	0,4114
kejut	0,4774	0	0
klaim	0	0	0,3484
litbang	0	0,3609	0
menang	0	0,2634	0
pacitan	0,4774	0	0
pdip	0	0	0,1956
pemilu	0	0	0,2328
prabowo	0	0,2171	0
proporsional	0	0	0,3586
salip	0	0,4028	0
santai	0	0,3783	0
sby	0,357	0	0
sebentar	0	0,3783	0
sistem	0	0	0,3277
soal	0,2647	0	0
survei	0	0,2809	0
temu	0,2661	0	0

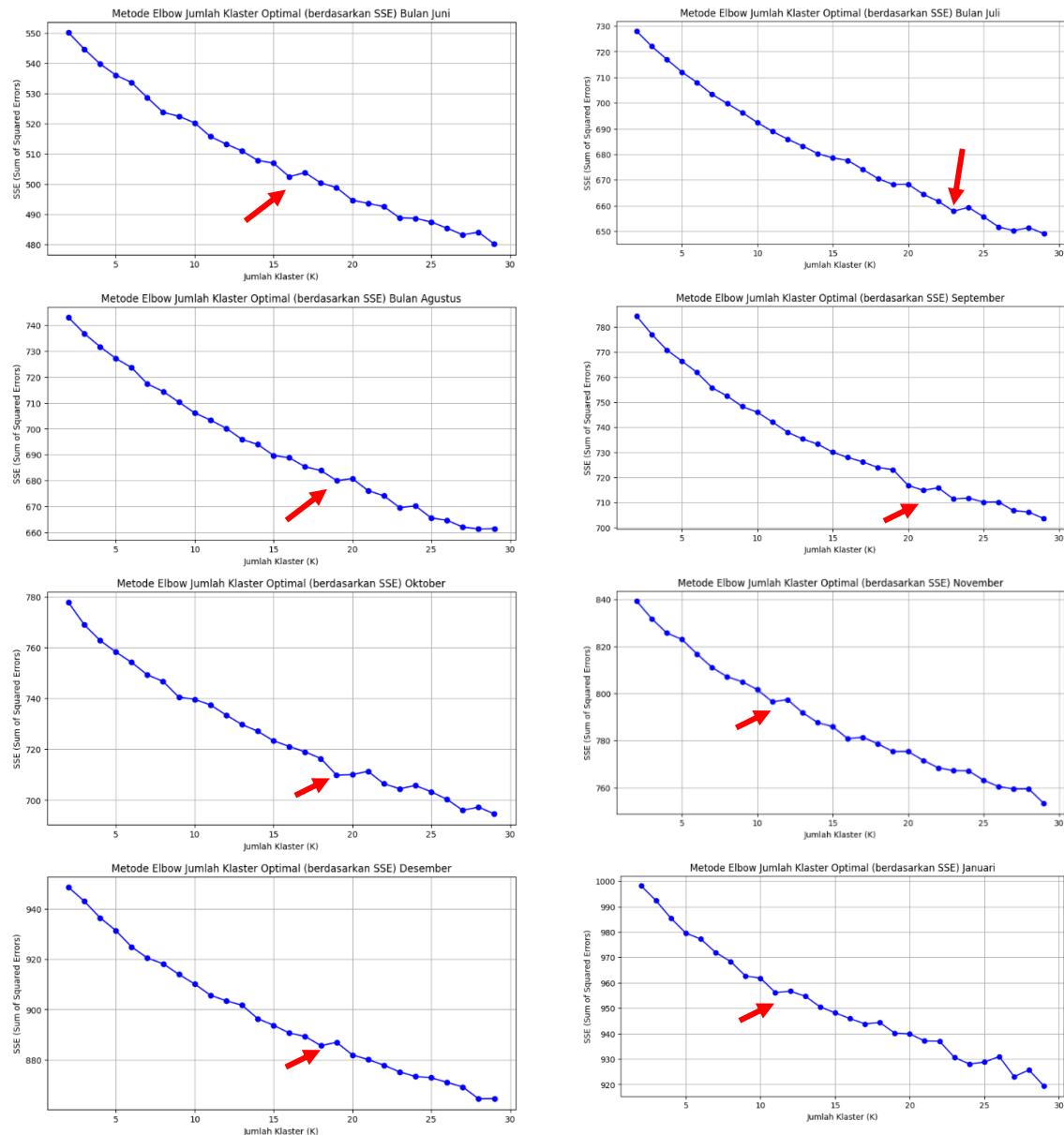
3.3. Klasterisasi

Tahapan klasterisasi diawali dengan penentuan jumlah klaster. Kemudian dilanjutkan dengan proses klasterisasi dengan algoritma K-Means. Terakhir adalah evaluasi dengan menggunakan DBI.

Gambar 3 terdapat delapan grafik *elbow method* dari Dataset bulan Juni 2023 hingga Januari 2023 untuk berita Kompas. Grafik pertama (baris pertama sebelah kiri) merupakan *elbow* yang terbentuk untuk klaster optimal bulan Juni yaitu dengan nilai $k=16$. Grafik kedua (baris pertama

sebalah kanan) merupakan *elbow* yang terbentuk untuk kluster optimal bulan Juli yaitu dengan nilai $k=23$. Grafik ketiga (baris kedua sebelah kiri) merupakan *elbow* yang terbentuk untuk kluster optimal bulan Agustus yaitu dengan nilai $k=19$. Grafik keempat (baris kedua sebelah kanan) merupakan *elbow* yang terbentuk untuk kluster optimal bulan September yaitu dengan nilai $k=21$. Grafik kelima (baris ketiga sebelah kiri) merupakan *elbow* yang terbentuk untuk kluster

optimal bulan Oktober yaitu dengan nilai $k=19$. Grafik keenam (baris ketiga sebelah kanan) merupakan *elbow* yang terbentuk untuk kluster optimal bulan November yaitu dengan nilai $k=11$. Grafik ketujuh (baris keempat sebelah kiri) merupakan *elbow* yang terbentuk untuk kluster optimal bulan Desember yaitu dengan nilai $k=18$. Grafik kedelapan (baris keempat sebelah kanan) merupakan *elbow* yang terbentuk untuk kluster optimal bulan Januari yaitu dengan nilai $k=11$.



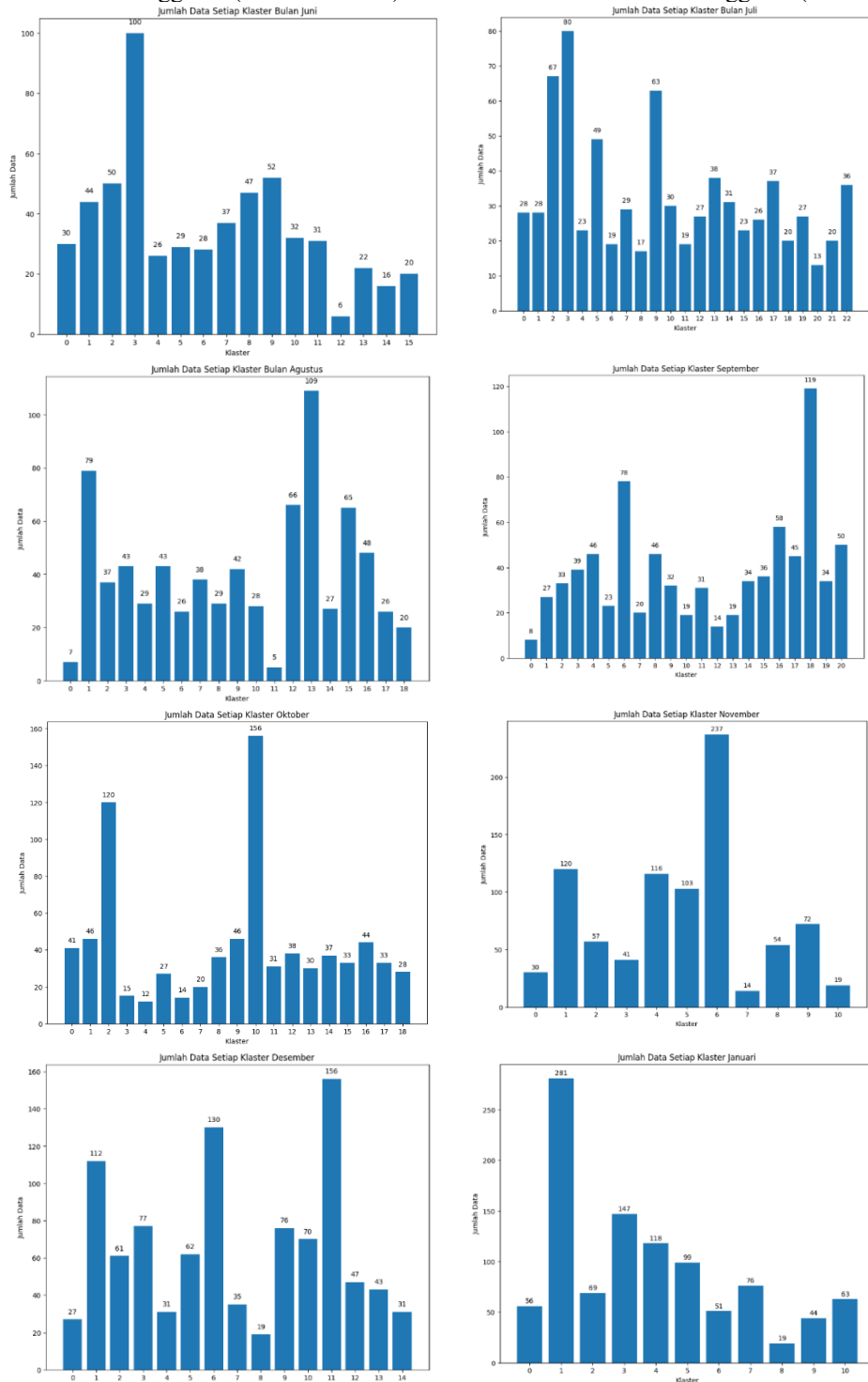
Gambar 3. Grafik Elbow Method Berita Kompas

Setelah mendapatkan kluster optimal berdasarkan *elbow* method maka setiap dataset dilakukan klasterisasi dengan nilai k optimal. Gambar 4 terdiri dari delapan gambar dari hasil jumlah data setiap kluster dari Dataset Berita Kompas. Grafik pertama (baris pertama sebelah kiri) merupakan jumlah data setiap kluster pada bulan Juni dari kluster 0 hingga 15 (total 16 kluster).

Grafik kedua (baris pertama sebelah kanan) merupakan jumlah data setiap kluster pada bulan Juli dari kluster 0 hingga 22 (total 23 kluster). Grafik ketiga (baris kedua sebelah kiri) merupakan jumlah data setiap kluster pada bulan Agustus dari kluster 0 hingga 18 (total 19 kluster). Grafik keempat (baris kedua sebelah kanan) merupakan jumlah data setiap kluster pada bulan September dari kluster 0 hingga

20 (total 21 klaster). Grafik kelima (baris ketiga sebelah kiri) merupakan jumlah data setiap klaster pada bulan Oktober dari klaster 0 hingga 18 (total 19 klaster). Grafik keenam (baris ketiga sebelah kanan) merupakan jumlah data setiap klaster pada bulan November dari klaster 0 hingga 10 (total 11 klaster).

Grafik ketujuh (beris keempat sebelah kiri) merupakan jumlah data setiap klaster pada bulan Desember dari klaster 0 hingga 14 (total 15 klaster). Grafik kedelapan (baris keempat sebelah kanan) merupakan jumlah data setiap klaster pada bulan Januari dari klaster 0 hingga 10 (total 11 klaster).



Gambar 4. Jumlah Data Setiap Klaster Berita Kompas

Evaluasi hasil klasterisasi menggunakan *Davies-Bouldin Index (DBI)* sebagai langkah untuk mengukur kualitas dan kecocokan klaster yang dihasilkan oleh algoritma *K-Means*. Tabel 9

merupakan hasil evaluasi setiap klaster. Evaluasi DBI menunjukkan bahwa setiap klaster optimal berdasarkan *elbow method* memiliki nilai DBI terkecil.

3.4. Penyajian Data

Penyajian data adalah dengan memilih klaster yang memiliki jumlah data terbanyak di setiap bulannya. Kemudian klaster terpilih tersebut dilakukan analisis dimensi dan elemen *framing*. Berdasarkan Gambar 4 dapat dilihat klaster dengan jumlah terbanyak di setiap bulannya.

Tabel 10 terlihat klaster dengan jumlah berita terbanyak dari Bulan Juni 2023 hingga Januari 2024 berdasarkan hasil klasterisasi berita Kompas.

Tabel 9. Nilai Davies-Bouldin Index (DBI) Berita Kompas

No.	Dataset	k	DBI
1	Berita Kompas Juni 2023	16	4,99
2	Berita Kompas Juli 2023	23	4,90
3	Berita Kompas Agustus 2023	19	5,27
4	Berita Kompas September 2023	21	5,28
5	Berita Kompas Oktober 2023	19	5,21
6	Berita Kompas November 2023	18	5,57
7	Berita Kompas Desember 2023	18	5,9
8	Berita Kompas Januari 2024	11	7,05

Tabel 10. Klaster dengan Jumlah Berita Terbanyak

No.	Bulan	Klaster	Jumlah Berita
1	Juni 2023	3	100
2	Juli 2023	3	80
3	Agustus 2023	13	109
4	September 2023	18	119
5	Oktober 2023	10	156
6	November 2023	6	237
7	Desember 2023	11	156
8	Januari 2024	1	281

Setelah dipilih klaster dengan jumlah berita terbanyak, peneliti selanjutnya melakukan analisis dimensi dan elemen *framing*. Dimensi *framing* dengan melakukan seleksi isu dan penonjolan aspek. Elemen *framing* dengan melakukan analisis *define problem*, *diagnose cause*, *make moral judgement* dan *treatment recommendation*. Analisis dimensi dan elemen framing melibatkan unsur interpretatif yang dipengaruhi subjektivitas dari peneliti. Oleh karena itu, analisis ini dilakukan secara sistematis menggunakan kerangka dari Entman dengan melibatkan beberapa pakar.

Tabel 11 merupakan hasil dari tahap pertama dimensi *framing* yaitu seleksi isu berita. Proses seleksi isu diperoleh dari hasil klasterisasi berupa frekuensi kata kunci dan koherensi topik tiap klaster. Dari setiap klaster dipilih isu yang menonjol berdasarkan jumlah berita dan analisis terhadap kata yang sering muncul secara konsisten. Dalam hal menjaga validitas interpretasi dimensi framing dilakukan oleh pakar komunikasi politik.

Tabel 11 menjelaskan dinamika isu yang menjadi sorotan utama dalam konteks politik dan pemilu dari Juni 2023 hingga Januari 2024. Pada bulan Juni, perhatian terfokus pada ketidakpastian pemilu, dinamika politik partai, dan persaingan antar calon presiden, yang kemudian berkembang pada Juli dengan isu seleksi calon wakil presiden dan komunikasi antar tokoh politik. Agustus menyoroti pemilihan daerah, hubungan politik, dan partai buruh, sementara September menekankan dinamika

calon wakil presiden dan dukungan partai. Pada Oktober, isu baru seperti kecerdasan buatan dan etika teknologi dalam politik muncul, dilanjutkan pada November dengan kritik terhadap sistem, generasi muda, dan erosi demokrasi. Desember menyoroti netralitas pejabat, berita *hoax*, dan persatuan bangsa, sedangkan

Januari 2024 menutup rangkaian isu dengan fokus pada kampanye pemilu, keterlibatan figur publik, serta kontroversi sosial-ekonomi dan politik.

Tabel 11. Seleksi Isu Berita Politik Kompas

No.	Bulan	Seleksi Isu
1	Juni 2023	a. Ketidakpastian pemilu b. Dinamika politik partai-partai besar c. Persaingan elektabilitas antar calon presiden d. Koalisi dan hubungan antar partai e. Pentingnya suara generasi muda f. Dinamika dalam tubuh internal partai g. Tantangan dan persiapan penyelenggaraan pemilu
2	Juli 2023	a. Seleksi Calon Wakil Presiden (Cawapres) b. Pertemuan dan Komunikasi Antar Tokoh Politik: c. Dinamika Internal Partai Politik d. Tantangan dan Persaingan dalam Koalisi Perubahan
3	Agustus 2023	a. Calon Wakil Presiden (Cawapres) b. Partai Buruh dan Prabowo c. Pemilihan Daerah dan Caleg d. Politik dan Partai Politik e. Hubungan dan Komunikasi Politik
4	September 2023	a. Rebutan Cawapres b. Gagasan duet Anies Baswedan dan Cak Imin c. Pendukung dan dukungan partai d. Dinamika internal partai demokrat e. Opsi Cawapres Prabowo
5	Oktober 2023	a. Kecerdasan Buatan dalam Pemilihan Pemimpin b. Implikasi Etika dan Keamanan Data: c. Pembatasan dan Regulasi d. Penerimaan Masyarakat Terhadap Penggunaan Teknologi Baru Dalam Proses Politik
6	November 2023	a. Panggung Politik Pilpres 2024 b. Kritik Sistem dan Korupsi c. Generasi Muda dan Kampanye Kreatif: d. Isu Ekonomi dan Kesenjangan: e. Erosi Demokrasi dan Etika Politik f. Peran Relawan dan Dukungan Tokoh g. Dinamika Hubungan Antara Partai dan Kandidat
7	Desember 2023	a. Netralitas Pejabat Polri b. Dukungan Partai Politik c. Peringatan terkait berita <i>Hoax</i> d. Persatuan Bangsa
8	Januari 2024	a. Dinamika Kampanye Pemilihan Presiden (Pilpres 2024) b. Kepentingan Politik dan Isu Integritas Tokoh c. Keterlibatan Figur Publik dalam Politik d. Kontroversi dan Kritik dalam Pemilu e. Isu Sosial-Ekonomi f. Kontroversi Pernyataan Tokoh

Setelah melakukan seleksi isu maka tahap kedua dan terakhir dari dimensi *framing* adalah

penonjolan aspek. Penonjolan aspek dianalisis berdasarkan pemakaian kata, kalimat, gambar dan citra tertentu yang ditampilkan media untuk konsumsi khalayak.

Tabel 12 merupakan *sample* dari penonjolan aspek berita berdasarkan salah satu seleksi isu berita di setiap bulannya dari pemberitaan Kompas tentang pemilu.

Tabel 12. Penonjolan Aspek Berita Kompas

Tanggal, Judul Berita dan	Penonjolan Aspek Gambar Tajuk	Penonjolan Aspek Kata/Kalimat
8 Juni 2023, "Ganjar Ingatkan Pentingnya Raih Suara Gen Z dan Milenial untuk Menangi Pemilu 2024"		"Menurut Ganjar Pranowo, kedua generasi itu merupakan pemilih potensial yang harus dimenangi oleh bacaleg PPP lantaran mewakili 50 persen jumlah pemilih."
1 Juli 2023, Soal Bakal Cawapres Anies, PKS: Anies Pulang, Kita Rapikan dan Umumkan.		"Terkait nama Cawapres, tim delapan KPP sepenuhnya telah menyerahkan pemilihan nama cawapres ke mantan Gubernur DKI Jakarta tersebut."

Pada Tabel 12 , penonjolan aspek dilakukan berdasarkan analisis terhadap berita yang telah terklasterisasi dan terpilih pada tahapan seleksi isu. Proses ini meliputi identifikasi elemen visual, dan bahasa yang digunakan oleh media, pemilihan judul berita, penggunaan gambar tajuk, dan penggunaan kalimat dalam isi berita. Bulan Juni adalah isu terkait pentingnya suara generasi muda, Bulan Juli adalah berita dari penonjolan aspek isu terkait calon wakil presiden. Bulan Agustus adalah berita dari penonjolan aspek isu terkait pemilihan daerah dan caleg. Bulan September adalah berita dari penonjolan aspek isu terkait pendukung dan dukungan partai. Bulan Oktober adalah berita dari penonjolan aspek isu terkait kecerdasan buatan dalam pemilihan pimpinan. Bulan November adalah berita dari penonjolan aspek isu terkait dinamika hubungan antara partai dan kandidat. Bulan Desember adalah berita dari penonjolan aspek isu terkait peringatan berita *hoax*. Bulan Desember adalah berita dari penonjolan aspek isu terkait kepentingan politik dan isu integritas tokoh. Penonjolan aspek dilihat berdasarkan judul berita, gambar tajuk, dan kata atau kalimat di dalam isi berita. Setelah dimensi *framing* selesai dilakukan, maka selanjutnya dilakukan analisis elemen *framing* berdasarkan berita yang diperoleh dari dimensi *framing*.

Tabel 13 menyajikan analisis terhadap berbagai berita terkait pemilu di Indonesia dengan pendekatan *framing* yang meliputi empat dimensi: *Define Problem*, *Diagnose Cause*, *Make Moral Judgment*, dan *Treatment Recommendation*. Setiap dimensi *framing* diperoleh dari menganalisis isi berita. Analisis dimensi dilakukan dengan mengidentifikasi teks berita yang menunjukkan masalah utama (*Define Problems*), menjelaskan penyebab masalah (*Diagnose Cause*), memberikan penilaian moral (*Moral Judgment*), dan menawarkan solusi atau rekomendasi tindakan (*Treatment Recommendation*).

Tabel 13. Elemen Framing Berita Kompas

Judul Berita	Define Problem	Diagnose Cause	Make Moral Judgment	Treatment Recommendation
"Ganjar Ingatkan Pentingnya Raih Suara Gen Z dan Milenial untuk Menangi Pemilu 2024"	Upaya memenangkan Pemilu 2024 dengan mendapatkan dukungan dari generasi milenial yang menyukai politik dan generasi Z, yang mewakili lebih dari 50% pemilih.	karakteristik khusus dari generasi Z dan milenial, yang tidak menyukai polemik politik dan lebih suka dialog yang egaliter dan toleran.	menarik dukungan dari generasi muda dianggap sebagai langkah positif dalam proses politik.	para calon legislatif dari PPP menggunakan kampanye yang inovatif dan kreatif untuk mendapatkan dukungan dari generasi milenial dan generasi Z.

Setiap berita membahas isu utama yang berbeda, seperti strategi meraih suara generasi muda, dinamika pemilihan calon wakil presiden, kontroversi *presidential threshold*, hingga wacana implementasi e-voting. Tabel 13 memberikan gambaran komprehensif tentang cara berita mengemas berita pemilu untuk memengaruhi persepsi publik, menawarkan perspektif atas dilema pemilu, dan merekomendasikan langkah-langkah strategis untuk penyelesaian masalah.

5. KESIMPULAN

Penelitian ini berhasil memanfaatkan algoritma *K-Means* untuk mengelompokkan berita pemilu di Kompas.com berdasarkan pola teks, yang kemudian digunakan untuk analisis *framing*. Temuan menunjukkan bahwa media secara aktif memilih dan menonjolkan aspek tertentu untuk membentuk persepsi publik terhadap isu politik. Hasil klasterisasi menunjukkan distribusi isu yang mencerminkan dinamika politik, seperti persaingan kandidat, isu koalisi, dan kebijakan regulasi pemilu. Elemen *framing* memberikan perspektif yang komprehensif tentang bagaimana media mendefinisikan masalah, mendiagnosis penyebab,

memberikan penilaian moral, dan merekomendasikan solusi.

Penelitian selanjutnya, disarankan untuk memperluas data berita dengan mencakup lebih banyak platform media dan periode waktu yang lebih panjang untuk mendapatkan hasil yang lebih representatif. Selain itu, mengintegrasikan analisis sentimen atau menggunakan algoritma klusterisasi yang lebih canggih dapat meningkatkan akurasi dan memberikan wawasan yang lebih mendalam. Penggunaan teknik visualisasi interaktif juga direkomendasikan untuk mempermudah interpretasi hasil analisis framing.

UCAPAN TERIMA KASIH

Terima kasih kepada Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi yang telah memberikan dukungan pendanaan melalui program Penelitian Dosen Pemula Tahun Anggaran 2024 atas terselenggaranya penelitian ini

DAFTAR PUSTAKA

- AFIDA, D., UDAYANTI, E. D., & KARTIKADHARMA, E. 2021. Aplikasi Text Mining untuk Klusterisasi Aduan Masyarakat Kota Semarang Menggunakan Algoritma K-means. *TRANSFORMATIKA*, 18(2), 215–224. <https://doi.org/http://dx.doi.org/10.26623/transformatika.v18i2.2362>
- AGUSTIANI, R., MUH UMRAN, L., & IBA, L. 2020. Analisis Framing Berita Politik Pemilihan Presiden Tahun 2019 pada Media Online Zonasultra.com. *Convergence: Jurnal Online Jurnalistik*, 2(1), 52–63. <http://ojs.uho.ac.id/index.php/Jurnalistik/index> 52
- ANITA, A., & ALIANDO, J. 2022. Pengelompokan Berita Kesehatan Pada Sosial Media Twitter Dengan Metode K-Means Clustering. *Ensiklopedia of Journal*, 4(3), 116–124. <https://doi.org/https://doi.org/10.33559/eoj.v4i3.877>
- AYUB, M. A. 2021. Analisis Topik Ekonomi dengan Algoritma K-Means Pada Media Online Era Pandemi COVID-19 di Sulawesi Tenggara. *Jurnal Informatika dan Komputer (JIKO)*, 4(2), 133–138. <https://doi.org/10.33387/jiko>
- BOER, K. M., PRATIWI, M. R., & MUNA, N. 2020. Analisis Framing Pemberitaan Generasi Milenial dan Pemerintah Terkait Covid-19 di Media Online. *Communicatus: Jurnal Ilmu komunikasi*, 4(1), 85–104. <https://doi.org/10.15575/cjik.v4i1.8277>
- FIRMAN, S., DESENA, W., & WIBOWO, A. 2022. Penerapan Algoritma Stemming Nazief & Adriani Pada Proses Klusterisasi Berita Berdasarkan Tematik Pada Laman (Web) Direktorat Jenderal HAM Menggunakan Rapidminer. *Syntax: Jurnal Informatika*, 11(2), 10–21. <https://doi.org/https://doi.org/10.35706/syji.v11i02.7192>
- HASAN, I., & FATHAN, F. 2020. Analisis Framing Berita Politik Menjelang Pilkada Klaten Tahun 2020 pada Solopos.com. *Academic Journal of Da'wa and Communication*, 1(2), 479–500.
- HIDAYAH, S. M., & RIAUAN, M. A. I. 2021. Analisis Framing Kebijakan Pembelajaran Tatap Muka di Media CNN Indonesia. *Medium*, 6(2), 167–184. [https://doi.org/https://doi.org/10.25299/medium.2021.vol9\(2\).8519](https://doi.org/https://doi.org/10.25299/medium.2021.vol9(2).8519)
- JO, T. 2019. *Text Mining Concept, Implementation, and Big Data Challenge*. Springer. <https://doi.org/https://doi.org/10.1007/978-3-319-91815-0>
- PURNIAWAN, I. M. A., SASMITA, G. M. A., & PRATAMA, I. P. A. E. 2022. Clustering Berita Menggunakan Algoritma TF-IDF dan K-Means dengan Memanfaatkan Sumber Data Crawling Pada Situs Detik.Com. *Jurnal Ilmiah Teknologi dan Komputer (JITTER)*, 3(1).
- RHEZA, M. A., & METANDI, F. 2020. Implementasi Metode K-Means Clustering untuk Penentuan Jenis Komentar Pada Tweet PSSI. *Jurnal Sains Terapan Teknologi Informasi (JUST TI)*, 12(2), 73–78. <https://doi.org/10.46964/justi.v12i2.363>
- SAHRIA YOGA. 2020. Implementasi Teknik Web Scraping pada Jurnal SINTA untuk Analisis Topik Penelitian Kesehatan Indonesia. *University Research Colloquium*, 297–306. <http://sinta2.ristekdikti.go.id/journals/detail>
- SIMANJUNTAK, H. T. A., SILABAN, P. E. P., MANURUNG, J. K. S., & SORMIN, V. H. 2023. Klusterisasi Berita Bahasa Indonesia Dengan Menggunakan K-Means Dan Word Embedding. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10(3), 641–652. <https://doi.org/10.25126/jtiik.20231026468>
- SIMILARWEB. 2023, December 5. *Most Visited News & Media Publisher Website ini Indonesia*. Similarweb.
- SIREGAR, A. K., & QURNIAWATI, E. F. 2022. Analisis Framing Pemberitaan Buzzer di Tempo.co. *Journal of New Media and Communication*, 1(1), 1–15.
- WIDANINGRUM, I., MUSTIKASARI, D., ARIFIN, R., TSAQILA, S. L., & FATMAWATI, D. 2022. Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan K-Means Clustering Untuk Menentukan Kategori Dokumen. *Prosiding Seminar Nasional Sistem Informasi dan Teknologi (SISFOTEK)*, 145–149.

YAZMI, R., MALA, A., HASNAN AFISAH, F.,
RADIKA SARI, F., HIKAM, I., RENDY
MAULANA, M., RAHIMA ORVALA, A., &
ILMU POLITIK, P. 2024. Diskursus Politik
Media Massa Indonesia: Analisis Framing

Berita Pemilu 2024 dalam Surat Kabar
Nasional Kompas. *Jurnal Majemuk*, 3(4),
704–715.
[http://jurnalilmiah.org/journal/index.php/maje
muk](http://jurnalilmiah.org/journal/index.php/majemuk)

Halaman ini sengaja dikosongi