student89 turnitin

ICAR 2025 Full Paper_AtikAriesta

Paper Malaysia

Document Details

Submission ID
trn:oid:::3618:106685069

Submission Date
Aug 1, 2025, 11:03 AM GMT+7

Download Date
May 26, 2026, 5:31 PM GMT+7

File Name
ICAR 2025 Full Paper_AtikAriesta.docx

File Size
3.4 MB

15 Pages

5,396 Words

31,361 Characters

22% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Match Groups

68 Not Cited or Quoted 18%
Matches with neither in-text citation nor quotation marks

10 Missing Quotations 2%
Matches that are still very similar to source material

3 Missing Citation 2%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

18% Internet sources
11% Publications
15% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **68** Not Cited or Quoted 18%
  Matches with neither in-text citation nor quotation marks
- **10** Missing Quotations 2%
  Matches that are still very similar to source material
- **3** Missing Citation 2%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

18% Internet sources
11% Publications
15% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
www.kompas.com		1%
2	Internet	
rmol.id		<1%
3	Internet	
jurnal.bppk.kemenkeu.go.id		<1%
4	Student papers	
University of Hull on 2024-09-03		<1%
5	Internet	
ejournal.undip.ac.id		<1%
6	Student papers	
Konsorsium Perguruan Tinggi Swasta Indonesia on 2025-05-05		<1%
7	Internet	
ejurnal.ung.ac.id		<1%
8	Internet	
journal.umy.ac.id		<1%
9	Internet	
e-jurnal.lppmunsera.org		<1%
10	Internet	
journal.formosapublisher.org		<1%

11 Internet
studenttheses.uu.nl <1%

12 Student papers
Erasmus University of Rotterdam on 2024-06-25 <1%

13 Student papers
University of Westminster on 2023-05-02 <1%

14 Internet
eprints.ums.ac.id <1%

15 Internet
www.ub.uni-mainz.de <1%

16 Student papers
Kaplan College on 2024-02-26 <1%

17 Internet
ijcsnet.id <1%

18 Internet
jurnal.umsu.ac.id <1%

19 Internet
ppp.or.id <1%

20 Student papers
Dublin City University on 2025-08-03 <1%

21 Internet
par.nsf.gov <1%

22 Publication
Christian Pierdzioch, Serkan Yener. "On the hump-shaped output effect of monet... <1%

23 Internet
hdl.handle.net <1%

24 Internet
repository.upiyptk.ac.id <1%

25 Student papers
Blue Mountain Hotel School on 2024-06-24 <1%

26 Internet
vbn.aau.dk <1%

27 Internet
www.repository.trisakti.ac.id <1%

28 Internet
www.teses.usp.br <1%

29 Internet
journal.budiluhur.ac.id <1%

30 Student papers
Taylor's Education Group on 2025-07-25 <1%

31 Student papers
Universiteit Utrecht on 2025-02-13 <1%

32 Student papers
University of Portsmouth on 2013-03-24 <1%

33 Student papers
Washoe County School District on 2021-07-03 <1%

34 Internet
docslib.org <1%

35 Student papers
Dublin City University on 2025-07-31 <1%

36 Student papers
Vrije Universiteit Brussel on 2025-05-23 <1%

37 Internet
discovery.researcher.life <1%

38 Internet
id.123dok.com <1%

39 Internet
repository.kln.ac.lk <1%

40 Internet
tecnoscientifica.com <1%

41 Student papers
University of South Wales - Pontypridd and Cardiff on 2025-06-06 <1%

42 Internet
hrmars.com <1%

43 Internet
idcorner.co.id <1%

44 Internet
journal.uinsgd.ac.id <1%

45 Internet
jurnalilmukomunikasi.uho.ac.id <1%

46 Student papers
Asia Pacific University College of Technology and Innovation (UCTI) on 2023-11-16 <1%

47 Student papers
Dublin City University on 2025-07-31 <1%

48 Publication
Farah Sherina Salma, Oktariani Nurul Pratiwi, Riska Yanu Farifah. "Classification ... <1%

49 Publication
M. Velasco-Juan, J. Fujioka. "Lagrangian nonlocal nonlinear Schrödinger equation... <1%

50 Student papers
Radboud Universiteit on 2025-07-16 <1%

51 Publication
Umair Ahmed, Muhammad Saeed, Shah Jamal Alam. "The role of suspicious acco... <1%

52 Internet
ebin.pub <1%

53 Internet
ejournal.unitomo.ac.id <1%

54 Internet
link.springer.com <1%

55 Internet
ojs.unpkediri.ac.id <1%

56 Internet
thesai.org <1%

57 Internet
www.jurnal.polgan.ac.id <1%

58 Internet
www.lingen.org.uk <1%

59 Publication
"SCSE 2022 Conference Proceedings", 2022 International Research Conference on ... <1%

60 Student papers
University of Winchester on 2024-08-02 <1%

61 Publication
Felipe Brubeck-Hernandez, Tanya Vladimirova, Mike Pooley, Robin Thompson, Br... <1%

International Conference on Cross- Disciplinary Academic Research 2025(ICAR 2025)
Online/Physical Conference
Paper No. XXX (The number assigned by the ICAR Committee)

Text Clustering as A Computational Approach for Analysing the Framing of Election News on Detik.com

Atik Ariesta
Universitas Budi Luhur
Jalan Ciledug Raya Petukangan Utara, Jakarta Selatan 12260, Jakarta, Indonesia
atik.ariesta@budiluhur.ac.id

ABSTRACT

This study develops a computational approach based on text clustering to analyse election news framing on the Detik.com portal. In the context of elections, media framing becomes an essential instrument in shaping public opinion towards political issues and actors. This study applies the K-Means algorithm to group news based on the similarity of text patterns in news titles, using TF-IDF weighting for feature representation. Data were collected through web scraping from the news.detik.com/pemilu channel between June 2023 and January 2024, resulting in more than 16,000 news items for analysis. This study also includes the development of a web application prototype using the Flask framework, which automates the text clustering process and supports framing analysis. This application offers features for exploring clustering results and word cloud visualization, enabling researchers to identify framing elements based on Entman's theory—specifically, defining problems, diagnosing causes, making moral judgments, and recommending treatments. Cluster quality evaluation was conducted using the Davies-Bouldin Index (DBI), which yielded consistent cluster results and warranted further analysis. The results of this study indicate that the clusters formed represent various news patterns, such as candidate support narratives, coalition dynamics, and election regulations. This study not only offers a quantitative approach to media framing analysis but also produces an intelligent system-based tool that can be replicated in other political text studies. The findings in this study broaden the understanding of the application of text mining techniques for media framing analysis in the realm of digital political communication.

KEYWORDS: text clustering, media framing, election, k-means, web application

1 INTRODUCTION

The media plays a crucial role in shaping public understanding of political issues, particularly during elections. Framing in news reporting is a method that emphasizes certain aspects of reality while downplaying others to influence how the public perceives an issue. Entman (1993) Framing theory identifies four core functions in media discourse: defining problems, diagnosing causes, making moral judgments, and suggesting solutions.

In the context of reporting on the 2024 Indonesian general election, Detik.com, as a leading online news outlet, has become a primary source of political information for the public. The media plays a crucial role in shaping public perceptions of candidates, political issues, and election results. Detik.com, as one of the most influential news sites in Indonesia, plays a crucial role in shaping public opinion. According to data from Similarweb, Detik.com ranks second in the News & Media Publisher category for website visits.

However, few studies have utilized automated approaches to detect framing in Indonesian online media coverage, particularly through clustering techniques for election headlines.

This study used an unsupervised learning approach, specifically the K-Means algorithm, because this algorithm is one of the most popular and easily understood clustering methods (Hasan et al., 2022; Ikotun et al., 2023). Previous research has shown that K-Means is effective in clustering Bangla-language news documents using the TF-IDF approach. Simanjuntak et al. (2023) also applied K-Means to analyse clusters of text data from social media, specifically Twitter. Their results showed dominant clusters related to economics, politics, social issues, and sports. The analysis was conducted using Orange Data Mining, including lexicon-based sentiment analysis. Disayiram & Rupasingha (2022) examined the application of clustering algorithms to group English-language news articles into five main categories: politics, sports, health, technology, and business. News articles were converted into vectors using the TF-IDF method, then clustered using three algorithms: Expectation-Maximization (EM), Simple K-Means, and Hierarchical Clustering. Evaluation using WEKA showed that the EM algorithm provided the highest accuracy of 88.5%. This study also concluded that news content provided more accurate clustering results than using news headlines alone. Hamami & Dahlan (2024) developed an online news segmentation system based on the K-Means algorithm combined with the Elbow method. This system groups news based on content similarity from the Detik news portal. The number of clusters changes daily following emerging topic trends, and the segmentation results are visualized through a web application, making it easier for users to access popular topics without having to read the entire news story.

Ashfaq et al. (2022) examined the use of framing in Pakistani media coverage of political protests spread on Twitter. They collected 498 tweets from three major media outlets (Dawn, The News, and Express Tribune) between October 20 and November 20, 2019. They conducted a quantitative framing analysis based on thematic and episodic categories, tweet types, and subject matter. Meanwhile, a study by Anggoro et al. (2023) analysed media framing of women's representation as Indonesian presidential candidates in 2024 using Robert Entman's model and a qualitative approach. The analysis was conducted on articles from Republika.com and Sindonews.com. The results show that Sindonews.com highlights the presence of female candidates positively, but is still limited to gender stereotypes as vice presidential candidates. At the same time, Republika.com places greater emphasis on electability and objective party support. Both media outlets contribute to shaping public perception of female candidates, despite bias in their reporting.

In general, framing analysis is still primarily conducted manually, which is time-consuming and tends to be subjective. Therefore, computational approaches such as text clustering are needed to provide a more efficient and objective analysis alternative. Furthermore, the development of a web-based application prototype that can automate news processing and facilitate systematic framing analysis is also needed.

2 METHODOLOGIES

2.1 Research Framework

This research adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, which consists of: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Implementation (Wirth & Hipp, 2000).

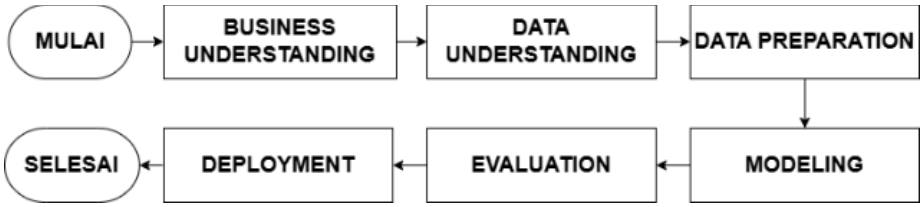

Figure 1: CRISP-DM Steps

The first step in this study is business understanding. In this step, the researcher determines the research objectives and identifies the problem to be solved. In this study, the problem identification is the clustering of election news for the analysis of election news framing on the election news website,

news.detik.com/pemilu. Framing is conducted to understand how the mass media actively selects and emphasizes certain aspects to shape public perception (Yazmi et al., 2024). This step will result in the determination of the main research problem related to the clustering of election news texts for framing analysis.

The second step is data understanding. Election news was collected from news.detik.com/pemilu using web scraping. The procedure for automatic extraction of data from websites using software (Khder, 2021). The tool used is a web scraper, an add-on for Google Chrome. The data to be collected covers news from June 2023 to January 2024. This timeframe was chosen because this period is the lead-up to the presidential and vice-presidential elections, which are characterized by escalating political issues, campaign strategies, and high media attention to the democratic process.

The third step is data preparation. Researchers conduct text processing, consisting of tokenization, stop word removal, and stemming (Jo, 2019). The text to be processed is a news headline. Tokenization involves removing special characters and numbers, converting to lowercase, and breaking the text based on whitespace. Stopword removal involves removing irrelevant words that frequently appear in the text, such as prepositions and conjunctions. Stemming is performed based on the results of stop word removal. Stemming converts each token containing an affix into its root form. The stemming process uses Sastrawi, the most appropriate stemmer for Indonesian (Firman et al., 2022). After data processing, researchers weigh words using Term Frequency-Inverse Document Frequency (TD-IDF). The output of the third step is a news headline that has undergone text processing and a TF-IDF weight matrix for each word in the headline.

The fourth step is modelling. Researchers cluster the headlines, which have already received their weighted values. Before carrying out the clustering process, the optimal k value is determined using the elbow method. The elbow shape of the elbow method graph is the k value that will be used to determine the cluster value. The output of this step is the optimal number of clusters determined using the elbow method, and news title clustering is carried out based on the TF-IDF value. The K-Means algorithm works by randomly initializing k cluster centres and then allocating each data item to a cluster based on its proximity to the nearest cluster centre (Widaningrum et al., 2022).

The fifth step is evaluation. Clustering results are evaluated using the Davies-Bouldin Index (DBI) to measure the quality and suitability of the clusters generated by the K-Means algorithm. The DBI for a given cluster is calculated as the ratio of Intra-Cluster Dispersion to Inter-Cluster Distance using Euclidean Distance (Anita & Aliando, 2022). The result of this step is the DBI value for each dataset.

The sixth step is deployment. In this step, researchers will select the cluster with the most significant amount of data each month. From the selected clusters, researchers determine the framing dimensions and framing elements. The framing dimensions include issue selection and aspect emphasis. The framing elements include defining the problem, diagnosing the cause, making moral judgments, and recommending treatment. Defining the problem is the element used to define the problem, relating how a research issue can be viewed and what the problem is. Diagnosing the cause is the element used to estimate the source of the problem, relating to the events seen as being caused by the events, and identifying the actors considered to be the cause of the problem. Making moral judgment is an element of moral decision-making that explains the problem or values used to delegitimize an action. Treatment recommendations are elements for solving the problem, including the proposed solutions to address the issue, and/or the proposed pathways to address the problem (Boer et al., 2020). The output of this step is the framing dimensions and framing elements in election news.

2.2 Prototype

A web-based prototype was built using the Flask framework. It includes features for text processing, text clustering, and framing data analysis. The prototype functions as a platform for examining framing patterns in political news.

3 RESULT AND DISCUSSIONS

3.1 Data Collection

Data shows a significant upward trend from June 2023 to January 2024, with the initial count of 1,194 cases nearly tripling to 3,111. A sharp spike occurred from September to December 2023, indicating rapid growth during that period. Overall, the total for the eight months reached 16,259.

Table 1. Number of news stories per month

No	Month Year	Amount
1	Juni 2023	1.194
2	Juli 2023	1.279
3	Agustus 2023	1.613
4	September 2023	1.910
5	Oktober 2023	2.308
6	November 2023	2.185
7	Desember 2023	2.659
8	Januari 2024	3.111
Total		**16.259**

Table 2 shows the attributes generated from web scraping on the Detik.com website, specifically news.detik.com/pemilu. Eight attributes were obtained using the Google Chrome web scraper tool.

Table 2. Attribute Scrapping

No	Attribute	Description
1	*web-scrapper-order*	The sequence of web scraper tools scraping news data
2	*web-scrapper-start-url*	The initial URL link of the web scraper scraping news data
3	Berita	Contains the date and title of the scraped news story
4	berita-href	The URL link of the scraped article/news story
5	Judul	The title of the scrapped article
6	Tanggal	The date of the scrapped article
7	GambarBerita-*src*	The main image link of the scraped news story
8	CaptionGambar	The short text of the main image scraped
9	IsiArtikel	The content of the scraped news story
10	Halaman	The page of the scrapped news story in the news list
11	Halaman-*href*	The URL link of the scraped news story

Table 3 shows the attributes determined for use in the clustering process: News Date, News Title, News Link, News Image Link, and News Content. The attributes were selected based on the needs of the news framing analysis. The clustering used the news title attribute because titles often contain keywords and the media's perspective, which serve as indicators of framing direction. Therefore, titles are highly relevant in assisting with the topic identification and framing elements in the analysis.

Table 3 Attributes Used

No.	Dataset Attribute	Description
1	Tanggal	News Date
2	Judul	News Title
3	Link Berita	News Link
4	Link Gambar	News Main Image Link
5	Isi Berita	News Content

Table 4 is the Detik news dataset for June 2023, which represents one of the eight datasets generated.

Table 4. Detik News Research Dataset for June 2023

Date	Title	News Link	Image Link	News Content
6/1/2023	Romahurmuziy: Sandiaga Uno Sudah 96 Persen Gabung PPP	https://news.detik.com/pemilu/d-6749932/romahurmuziy-sandiaga-uno-sudah-96-persen-gabung-ppp	https://akcdn.detik.net.id/community/media/visual/2023/06/01/romahurmuziy_169.jpeg?w=700&q=90	Jakarta - Ketua Majelis Pertimbangan PPP, Romahurmuziy menyebut Sandiaga Uno sudah tinggal menunggu waktu untuk bergabung dengan PPP. Romy pun mengatakan bahwa Sandi sudah 96 persen bergabung dengan PPP ……..
. . .	. . .	. . .	. . .	. . .
6/30/2023	Beredar SE Walkot Depok Soal Penertiban Spanduk hingga Atribut Partai	https://news.detik.com/pemilu/d-6799957/beredar-se-walkot-depok-soal-penertiban-spanduk-hingga-atribut-partai	https://akcdn.detik.net.id/community/media/visual/2023/06/23/walkot-depok-m-idris-devidetikcom_169.jpeg?w=700&q=90	Jakarta - Surat edaran (SE) Wali Kota Depok terkait penertiban pemasangan bendera, spanduk hingga atribut beredar di publik. SE tersebut ditujukan ke Ketua DPC atau DPD Partai Politik se-Kota Depok…….

3.2 Text Processing

Text processing in this study began with tokenization, stopword processing, and stemming. The attribute processed for use in the study was the title. Table 6 shows an example of the process of removing special characters and numbers, converting to lowercase, and tokenization (Jo, 2019).

Table 5. Transform and Tokenize News Titles

Initial Title	Transform Text	Tokenization
Kritik Jokowi, Puskapol UI: Dalih Cawe-cawe Pilpres untuk "Bangsa dan Negara" Alasan Klise	kritik jokowi puskapol ui dalih cawecawe pilpres untuk bangsa dan negara alasan klise	['kritik', 'jokowi', 'puskapol', 'ui', 'dalih', 'cawecawe', 'pilpres', 'untuk', 'bangsa', 'dan', 'negara', 'alasan', 'klise']

Table 5 shows the text processing of the initial title, which will then be transformed by removing special characters and numbers and converting them to lowercase. The Transform Text column displays the removal of special characters such as double quotation marks ("") and colons (:) and has been converted to all lowercase. The Tokenization column displays the tokenization results, which separate news titles based on whitespace.

Table 6 Stopword Removal News Title

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

Table 6 shows the results of tokenization, which will undergo stopword removal. As seen in the Stopword Removal column, the words "untuk," "cawecawe," and "dan," which are conjunctions, are no longer present. Stemming also combines tokens into a single sentence.

Table 7 Stemming News Title

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

Table 7 shows the stemming results from stop word removal and combining each token into a sentence. After stemming, the word "alasan" is replaced by its root word, "alas," as shown in the stemming column.

The next step in text processing is to weight the words in the news titles. Table 8 shows an example of TF-IDF results from three news items: News 1, News 2, and News 3. Cells with a number indicate the news item contains the word corresponding to the column, while cells with a zero indicate the news item does not contain the word corresponding to the column. News 1 contains the words "ada," "anies," "demokrat," "kejut," "pacitan," and "sby." News 2 contains the words "elekktablitas," "ganjar," "litbang," "menang," "prabowo," "salip," "santai," "sebentar," and "survei." News 3 has the words “Airlangga”, “Open”, “Caleg”, “Want”, “Claim”, “PDIP”, “Election”, “Proportional”, “System”.

Tabel 1. Term Frequency – Inverse Document Frequency (TF-IDF) News Title

	News 1	News 2	News 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 Clustering

The clustering process begins with determining the number of clusters. This is followed by the clustering process using the K-Means algorithm. Finally, evaluation is performed using the DBI algorithm.

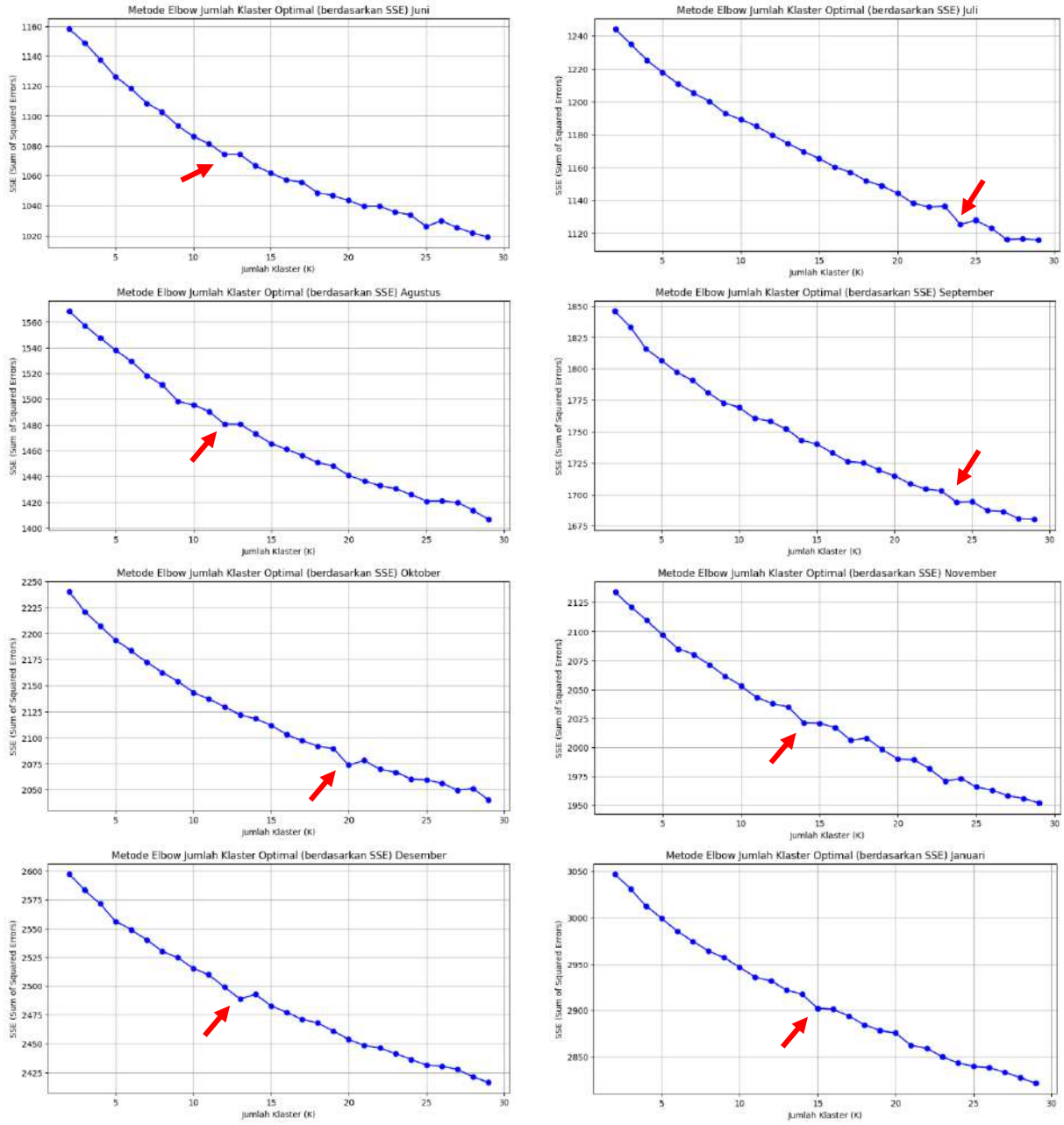

Figure 2: Elbow Method Chart Detik News

Figure 2 shows eight elbow method graphs from Datasets 9 to 16 for Detik news. The first graph (first row on the left) is the elbow formed for the optimal cluster in June with a value of k = 12. The second graph (first row on the right) is the elbow formed for the optimal cluster in July with a value of k = 24. The third graph (second row on the left) is the elbow formed for the optimal cluster in August with a value of k = 12. The fourth graph (second row on the right) is the elbow formed for the optimal cluster in September with a value of k = 24. The fifth graph (third row on the left) is the elbow formed for the optimal cluster in October with a value of k = 20. The sixth graph (third row on the right) is the elbow formed for the optimal

cluster in November with a value of k = 14. The seventh graph (fourth row on the left) is the elbow formed for the optimal cluster in December with a value of k = 13. The eighth graph (fourth row on the right) is the elbow formed for the optimal cluster in January, namely with a value of k = 15.

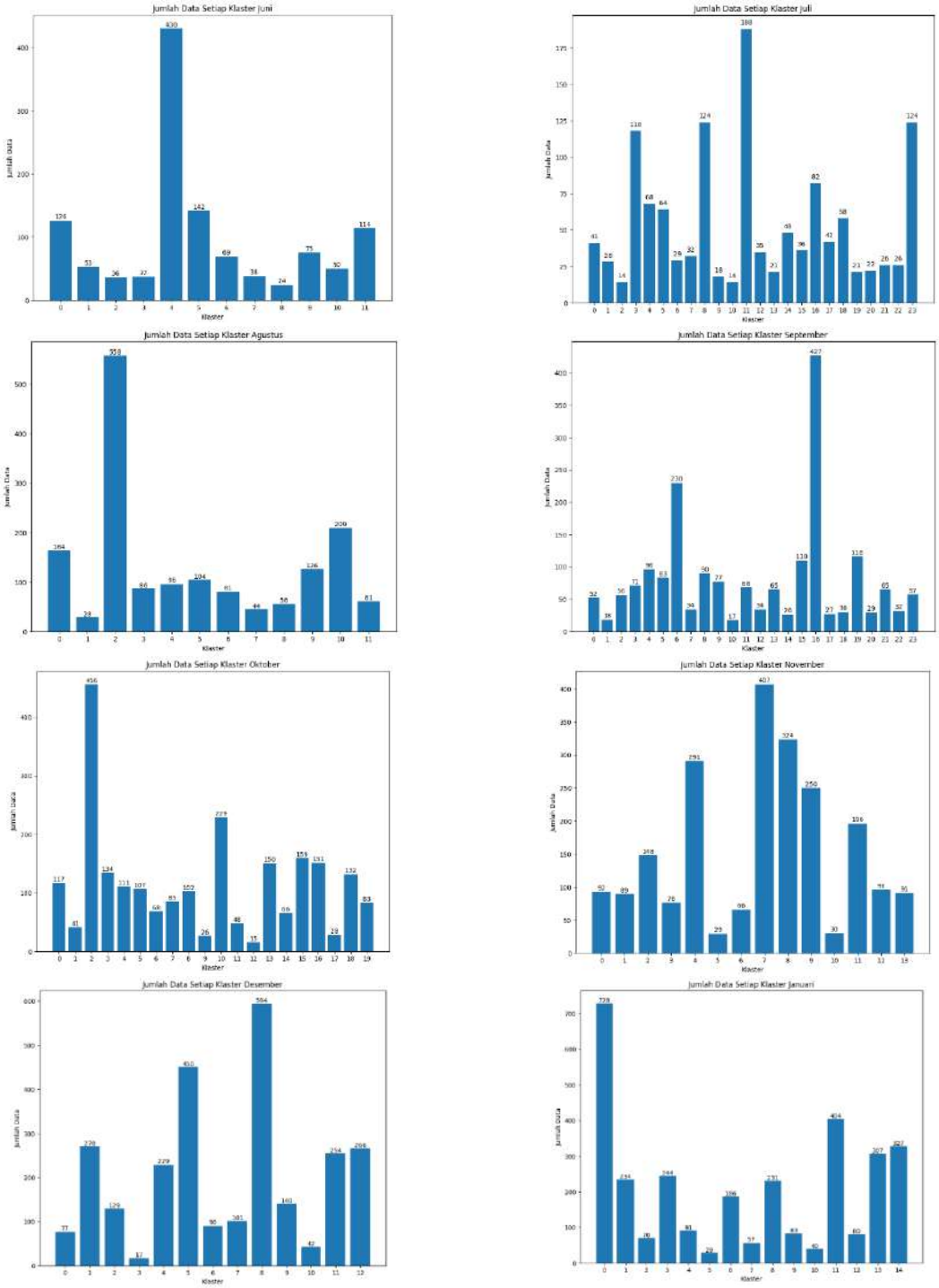

Figure 3: Number of Data for Each Detik News Cluster

After obtaining the optimal cluster based on the elbow method, each dataset was clustered with the optimal k value. Figure 3 consists of eight images of the results of the amount of data for each cluster from the Detik News Dataset. The first graph (first row on the left) is the amount of data for each cluster in June from cluster 0 to 11 (a total of 12 clusters). The second graph (first row on the right) is the amount of data

for each cluster in July from cluster 0 to 23 (a total of 24 clusters). The third graph (second row on the left) is the amount of data for each cluster in August, from cluster 0 to 11 (a total of 12 clusters). The fourth graph (second row on the right) shows the data for each cluster in September from cluster 0 to 24 (a total of 25 clusters). The fifth graph (third row on the left) shows the data for each cluster in October from cluster 0 to 189 (a total of 20 clusters). The sixth graph (third row on the right) shows the total data for each cluster in November, from clusters 0 to 13 (a total of 14 clusters). The seventh graph (fourth row on the left) shows the total data for each cluster in December, from clusters 0 to 12 (a total of 13 clusters). The eighth graph (fourth row on the right) shows the total data for each cluster in January, from clusters 0 to 14 (a total of 15 clusters).

The clustering results were evaluated using the Davies-Bouldin Index (DBI) to measure the quality and suitability of the clusters generated by the K-Means algorithm. Table 8 shows the evaluation results for each cluster. The DBI evaluation shows that each optimal cluster based on the elbow method has the smallest DBI value.

Table 8 Davies-Bouldin Index (DBI) Value Detik News

No.	Dataset	*k*	Davies-Bouldin Index (DBI)
1	News Detik Juni 2023	12	5,38
2	News Detik Juli 2023	24	5,12
3	News Detik Agustus 2023	12	6,04
4	News Detik September 2023	24	5,29
5	News Detik Oktober 2023	20	5,77
6	News Detik November 2023	14	6,41
7	News Detik Desember 2023	13	6,95
8	News Detik Januari 2024	15	6,72

3.4 Data Presentation

Data presentation was performed by selecting the cluster with the most significant amount of data each month. These selected clusters were then analysed for dimensions and framing elements. Figure 3 shows the clusters with the most significant number of news items for each month.

Table 10 shows the clusters with the most significant number of news items from June 2023 to January 2024, based on the Detik news clustering results.

Table 9 Cluster with the Most News

No.	Month	Cluster	News Amount
1	Juni 2023	4	430
2	Juli 2023	11	188
3	Agustus 2023	2	558
4	September 2023	16	427
5	Oktober 2023	2	456
6	November 2023	7	407
7	Desember 2023	8	594
8	Januari 2024	0	728

After selecting the cluster with the most significant number of news items, the researchers then analyzed the framing dimensions and elements. The framing dimensions involved selecting issues and highlighting aspects. The framing elements included defining the problem, diagnosing the cause, making moral judgments, and providing treatment recommendations. The analysis of the framing dimensions and elements involved interpretative elements influenced by the researcher's subjectivity. Therefore, this analysis was conducted systematically using Entman's framework, involving several experts.

Table 11 shows the results of the first stage of the framing dimensions, namely, news issue selection. The issue selection process was derived from clustering results, including keyword frequency and topic coherence for each cluster. From each cluster, prominent issues were selected based on the number of news items and analysis of consistently recurring words. To ensure the validity of the framing dimension interpretation, political communication experts conducted the analysis.

Table 10 Detik Political News Issue Selection

No.	Month	Selecting Issue
1	Juni 2023	a. Elections and National Political Dynamics b. Internal Dynamics of Political Parties c. Support from Volunteers and Community Leaders d. Social and Government Issues e. Surveys and Electability
2	Juli 2023	a. Coalitions and Political Strategy b. Gender and Representation c. National Policies and Issues d. Viral Phenomena and Public Response e. Regional Head Elections f. Social and Cultural Issues
3	Agustus 2023	a. Presidential and Vice-Presidential Candidates b. Coalition Dynamics and Party Relations c. Election Strategy and Campaign d. Internal Controversy and Criticism e. Policies and Amendments f. Regional Issues and Local Figures
4	September 2023	a. 2024 Presidential Election Issues b. Identity Issues and Political Polarization c. Maneuvers of Political Parties and Political Figures d. Social Dynamics and Public Support e. Special Issues and Supporters
5	Oktober 2023	a. The Political Dynamics of the 2024 Presidential and Vice Presidential Candidates b. The Battle of Narratives and Campaign Strategies c. External and Internal Party Conflicts d. Regulation and Neutrality in the Election e. Volunteer Visits and Declarations
6	November 2023	a. The Dynamics of Jokowi's Relationship with the Indonesian Democratic Party of Struggle (PDIP) b. Kaesang and the PSI's Political Maneuvers c. Election and Candidacy Context d. Internal and Cross-Party Conflict e. Megawati's Speeches and Statements
7	December 2023	a. Presidential Candidate Campaigns and Programs b. Elite Political Relations c. Social and Security Issues d. Surveys and Electability e. Regional-Specific Issues
8	Januari 2024	a. Political Campaigns and Promises b. Political Strategy and Elections c. Responses to National Issues d. Religion and Traditions e. Relationships with Other Figures and Parties

Table 10 presents a selection of issues for each month. The period from June 2023 to January 2024 marked significant developments in Indonesian politics leading up to the 2024 elections. Initially, the focus was on election preparations and internal party dynamics, then shifted to coalition formation and political strategy following the announcement of presidential and vice-presidential candidates. As the election approached, intensive discussions involved campaign strategies, narrative battles, and internal and cross-party conflicts. Political manoeuvres by key figures and the dynamics of elite relations dominated the news, including issues related to election regulations and neutrality. Finally, as election day approached, the focus shifted to campaign promises, final strategies, and responses to national issues, reflecting a crucial stage in the political contestation. January 2024 concluded the series with a focus on election campaigns, the involvement of public figures, and socio-economic and political controversies.

After selecting the issues, the second and final stage of the framing dimension is aspect emphasis. Aspect emphasis is analysed based on the use of specific words, sentences, images, and visuals presented by the media for audience consumption. Table 11 is a sample of news aspect emphasis based on a selection of news issues from each month of Detik's election coverage.

Table 11 Highlighting Aspects of Detik News

Date, News Title	Highlighting Aspects of Title Image	Emphasis on Word/Sentence Aspects
1 Juni 2023, "Ganjar Serukan Dukungan Partai dan Relawan Bersatu di Pilpres 2024"		"Ganjar menenkankan pentingnya berkolaborasi antara berbagai elemen dalam memenangkan pemilu. Dia menyerukan semua pihak bekerja sama demi mencapai kemenangan bersama. "Hari ini kita sedang menunjukkan ke publik bagaimana kolaborasi, kerja sama, bagaimana saling menghormati itu dilakukan karena kita punya cita-cita yang sama, menang," katanya."
2 Juni 2023, "Ketum PAN Bakal Datangi Markas PDIP Sore Ini, Bahas Dukungan ke Ganjar?"		"PAN telah menyerahkan urusan capres kepada Zulhas. "Jadi PAN besok malam juga ada rapat di DPP ya, salah satunya membahas isu kekinian termasuk masalah kontestasi Pilpres walaupun memang PAN itu sudah memberikan mandat penuh kepada Ketum," tutur Yandri. "Tapi update atau perkembangan kekinian perlu didiskusikan di internal partai dan kami, memang sampai sekarang masih melakukan komunikasi dengan semua pihak."

In Table 11, the aspect highlighting is based on an analysis of news items that have been clustered and selected during the issue selection stage. This process includes identifying visual elements and the language used by the media, selecting news headlines, using headline images, and using sentences in the news content.

Table 12 presents an analysis of various election-related news stories in Indonesia using a framing approach that encompasses four dimensions: Define Problem, Diagnose Cause, Make Moral Judgment, and Treatment Recommendation. Each framing dimension is derived from analyzing news content. Dimensional analysis is conducted by identifying news texts that address the main problem (Define Problems), explain the cause of the problem (Diagnose Cause), provide a moral judgment (Moral Judgment), and offer solutions or recommended actions (Treatment Recommendation).

Table 12 Detik News Framing Elements

News Title	Define Problem	Diagnose Cause	Make a Moral Judgment	Treatment Recommendation
Ganjar Serukan Dukungan Partai dan Relawan Bersatu di Pilpres 2024	Ganjar Pranowo menyerukan kolaborasi antara partai politik dan relawan untuk memenangkan Pilpres 2024. Masalah utama yang diidentifikasi adalah potensi ketidakharmonisasi atau kurangnya koordinasi antara partai politik dan relawan dalam mendukung calon presiden, yang bisa mengurangi efektivitas kampanye dan peluang kemenangan.	Meskipun partai politik memiliki struktur formal dan strategi yang terorganisasi, relawan biasanya bersifat independen dan lebih fleksibel, sehingga terkadang sulit untuk menyelaraskan tujuan dan cara kerja kedua pihak.	Seruan Ganjar untuk kolaborasi antara partai politik dan relawan adalah langkah positif yang menunjukkan semangat demokrasi gotong-royong. Namun, penting untuk memastikan bahwa kolaborasi ini tidak hanya bertujuan untuk kemenangan politik, tetapi juga mencerminkan nilai-nilai yang menghormati kedaulatan rakyat, keterbukaan, dan inklusivitas. Keseimbangan antara kepemimpinan partai dan aspirasi relawan perlu dijaga agar tidak mengalienasi salah satu pihak.	Bentuk tim penghubung antara partai politik dan relawan untuk menyelaraskan strategi kampanye lalu Pastikan bahwa semua pihak memahami dan mendukung visi serta misi calon secara konsisten. Kemudian yang terpenting berikan pelatihan kepada relawan untuk menyampaikan pesan kampanye dengan cara yang efektif tanpa mengesampingkan nilai-nilai demokrasi.

Table 13 provides a comprehensive overview of how news stories package election news to influence public perception, offers perspectives on election dilemmas, and recommends strategic steps for problem-solving.

3.5 Web Application Prototyping

The web application was developed using the Flask framework and deployed locally. This application has several modules for text processing, clustering, visualization, and framing analysis.

The framing analysis module inputs framing data such as problem definitions, causal diagnoses, moral judgments, and treatment recommendations, while simultaneously using clustering to enhance interpretability further. This module serves as a link between automatic clustering and framing analysis.

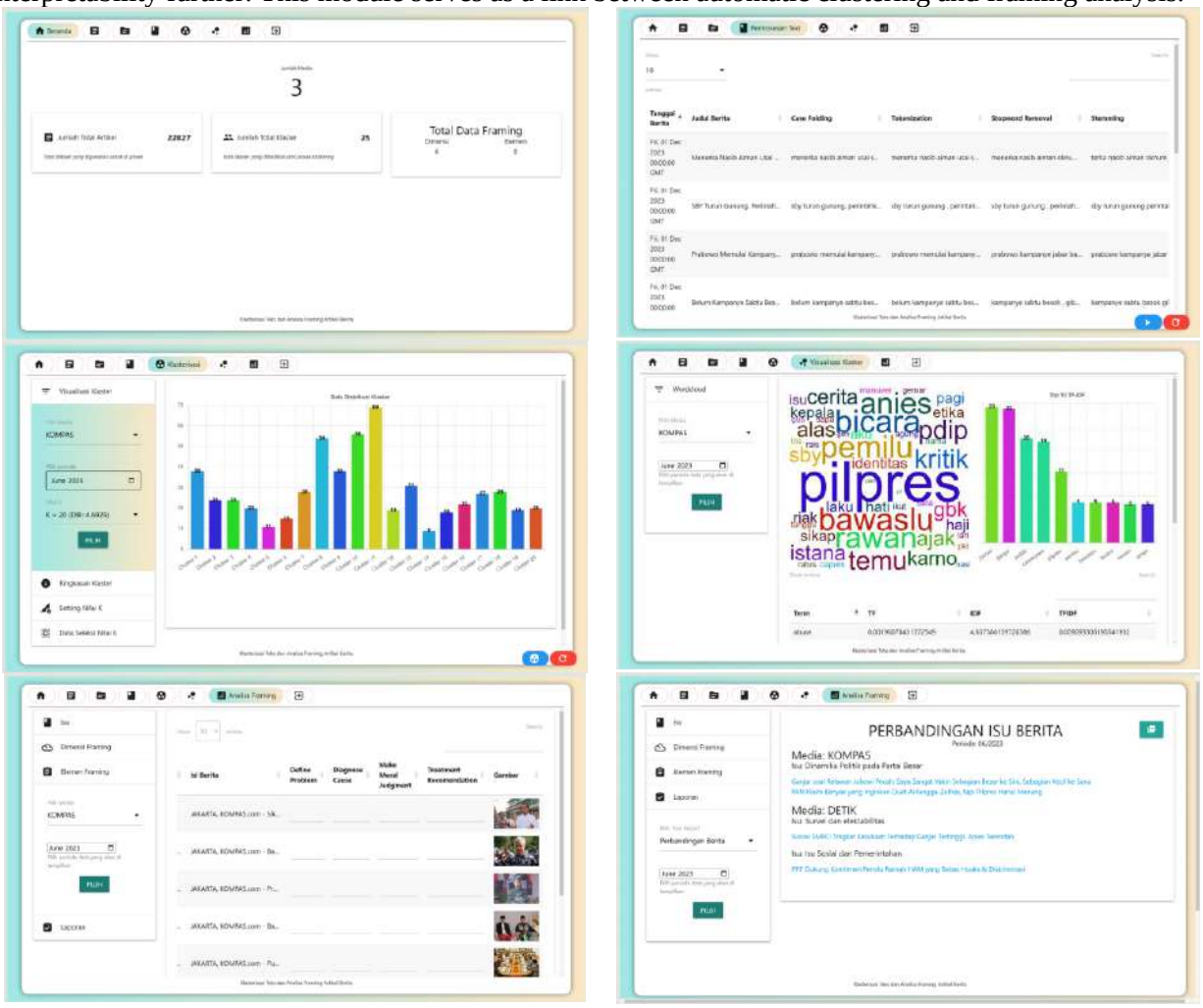

Figure 4: Application Screenshot

Figure 4 is a screenshot of the application interface to clarify the system's functionality. While the core analysis in this study focuses on Detik.com headlines, the user interface screenshots in this paper show test data sourced from Kompas.com for layout and visual validation purposes. Compared to manual framing analysis, this computational approach demonstrates speed and scalability, while maintaining interpretability through visualization and clustering validation.

4 CONCLUSIONS

This study successfully utilized the K-Means algorithm to cluster election news on Detik based on text patterns, which were then used for framing analysis. The findings indicate that the media actively select and highlight certain aspects to shape public perceptions of political issues. The clustering results show a distribution of issues that reflect political dynamics, such as candidate competition, coalition issues, and

election regulatory policies. The framing elements provide a comprehensive perspective on how the media defines problems, diagnoses causes, provides moral judgments, and recommends solutions.

Future research is recommended to expand the news data by covering more media platforms and a longer period to obtain more representative results. Furthermore, integrating sentiment analysis or using more sophisticated clustering algorithms can improve accuracy and provide deeper insights. The use of interactive visualization techniques is also recommended to facilitate interpretation of the framing analysis results.

5 ACKNOWLEDGEMENTS

Thank you to the Directorate of Research, Technology, and Community Service, Directorate General of Higher Education, Research, and Technology, Ministry of Education, Culture, Research, and Technology, for providing funding support through the 2024 Fiscal Year Beginner Lecturer Research program for the implementation of this research.

REFERENCES

Anggoro, A. D., Puspitasari, F. A., & Som, A. P. M. (2023). Robert Entman's Framing Analysis: Female Representation in 2024 Presidential Candidates on Republika.com and Sindonews.com. *Komunikator*, *15*(2), 211–224. https://doi.org/10.18196/jkm.19247

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

Ashfaq, M., Shahid, N.-A., & Zubair, J. (2022). New Media and Political Protest: Framing Analysis of News Content on Twitter. *Asian Journal of Comparative Politics*, *7*(4), 1177–1189. https://doi.org/10.1177/20578911211067317

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

Disayiram, N., & Rupasingha, R. A. H. M. (2022). A Comparative Study of Clustering English News Articles Using Clustering Algorithms. *2022 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, 108–113. https://doi.org/10.1109/SCSE56529.2022.9905210

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Firman, S., Desena, W., & Wibowo, A. (2022). Penerapan Algoritma Stemming Nazief & Adriani Pada Proses Klasterisasi 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

Hamami, F., & Dahlan, I. A. (2024). Analyzing News Clustering with K-Means Algorithm: A Segmentation Perspective. *2024 4th International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, 42–46. https://doi.org/10.1109/ICE3IS62977.2024.10775603

Hasan, S. Al, Ruiqin, W., & Hussain, M. G. (2022). Clustering Analysis of Bangla News Articles with TF-IDF & CV Using Mini-Batch K-Means and K-Means. *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 17–22. https://doi.org/10.1109/CyberneticsCom55287.2022.9865339

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means Clustering Algorithms: A Comprehensive Review, Variants Analysis, and Advances In the Era of Big Data. *Information Sciences*, *622*, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

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

Khder, M. (2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing and Its Applications*, *13*(3), 145–168. https://doi.org/10.15849/IJASCA.211128.11

Simanjuntak, H. T. A., Silaban, P. E. P., Manurung, J. K. S., & Sormin, V. H. (2023). Klasterisasi 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

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.

Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a Standard Process Model for Data Mining*. https://api.semanticscholar.org/CorpusID:1211505

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/majemuk