

Analisis Sentimen Publik Media Sosial Terhadap Pemecatan Sty Menggunakan Metode K-Medoids

Wawan Jhon Putra Ziraluo¹, Indra Kelana Jaya², Harlen Gilbert Simanullang³
Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Received, Agst 30, 2025

Revised, Sept 17, 2025

Accepted, Sept 25, 2025

Keywords:

sentimen,
shin tae yong,
K-medoids,

ABSTRAK

Reaksi publik yang beragam dapat mempengaruhi opini umum dan menciptakan polarisasi di kalangan penggemar sepak bola. Dalam konteks ini, penting untuk memahami bagaimana sentimen masyarakat terbentuk dan bagaimana penyebaran opini berkembang di berbagai platform media sosial. Analisis sentimen menjadi salah satu cara efektif untuk mengidentifikasi pola opini publik terkait pemecatan pelatih Timnas Shin Tae Yong. Dalam mengidentifikasi sentiment Masyarakat dalam mengekspresikan pendapat metode yang sesuai untuk digunakan adalah metode k-medoids. Berdasarkan hasil pengklasteran maka cluster negatif sebanyak 122 komentar, cluster netral sebanyak 39 dan cluster positif sebanyak 39.

This is an open access article under the [CC BY-SA](#) license.



Penulis Koresponden:

Wawan Jhon Putra Ziraluo,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: jhonwawan45@gmail.com

1. PENDAHULUAN

Sepak bola merupakan salah satu cabang olahraga dengan jumlah pendukung yang sangat besar di Indonesia[1]. Topik yang berkaitan dengan tim nasional kerap menjadi sorotan publik, terutama ketika menyangkut keputusan penting seperti pergantian pelatih. Salah satu peristiwa yang menimbulkan reaksi luas adalah pemberhentian pelatih tim nasional, Shin Tae-yong. Keputusan tersebut memicu beragam respons di media sosial—mulai dari dukungan hingga kritik—yang mencerminkan keberagaman persepsi masyarakat terhadap kebijakan yang diambil. Saat ini, media sosial berperan penting sebagai ruang publik virtual tempat masyarakat mengekspresikan pendapat secara cepat dan terbuka. Dalam konteks ini, analisis sentimen menjadi metode yang relevan untuk menangkap kecenderungan opini publik terhadap isu-isu tertentu. Melalui analisis komentar di berbagai platform seperti Twitter, Instagram, dan TikTok, peta persepsi masyarakat dapat dibentuk dan digunakan sebagai referensi dalam proses pengambilan keputusan, khususnya di bidang olahraga.

Untuk kebutuhan tersebut, penelitian ini menggunakan algoritma K-Medoids—sebuah teknik *clustering* berbasis partisi yang unggul dalam mengelola data dengan outlier dan distribusi yang tidak seragam[1]. Berbeda dengan K-Means yang menggunakan nilai rata-rata sebagai pusat klaster [2], K-Medoids memilih titik pusat yang benar-benar ada dalam data, sehingga lebih stabil dan cocok untuk analisis teks dari media sosial yang bersifat tidak terstruktur.

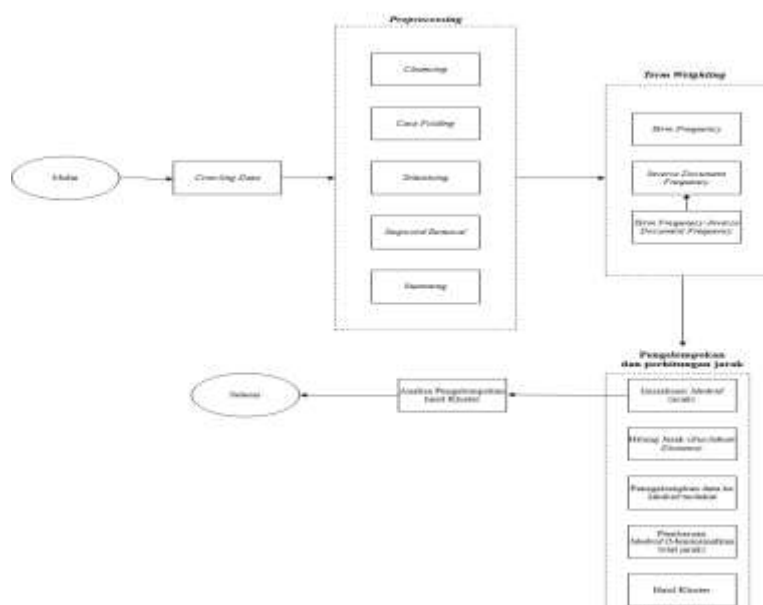
2. METODE PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif dengan metode *unsupervised learning*, di mana algoritma K-Medoids digunakan untuk mengelompokkan sentimen masyarakat terkait pemberhentian Shin Tae-yong. Sumber data berasal dari tiga media sosial utama—Twitter, Instagram, dan TikTok—yang dikumpulkan melalui teknik *web scraping* pada komentar-komentar publik setelah pengumuman resmi pemecatan.

Data yang terkumpul kemudian diproses melalui beberapa tahap pembersihan dan normalisasi teks, antara lain *cleansing*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*, untuk menghasilkan format teks yang seragam dan siap dianalisis[3]. Setelah pra-pemrosesan, teks dikonversi menjadi bentuk numerik menggunakan pendekatan Term Frequency–Inverse Document Frequency (TF-IDF), guna memberikan bobot pada setiap kata berdasarkan frekuensi dan penyebarannya[4].

Proses klusterisasi dilakukan menggunakan algoritma K-Medoids, yang membentuk kelompok berdasarkan kemiripan antar data dengan menghitung jarak menggunakan Euclidean Distance[5]. Pusat kluster (medoid) dipilih dari data aktual, dan proses ini dilakukan secara berulang hingga hasil klusterisasi stabil.

Kualitas kluster yang terbentuk dievaluasi menggunakan metrik *Silhouette Score*, yang menunjukkan tingkat konsistensi antar anggota kluster. Untuk memperjelas hasil analisis, visualisasi data juga disajikan, dan keseluruhan proses dituangkan dalam skema desain penelitian yang sistematis.



Gambar 1 Desain Penelitian

3. HASIL DAN PEMBAHASAN

Penelitian ini berhasil mengelompokkan sentimen publik terhadap pemecatan Shin Tae-yong berdasarkan data yang dikumpulkan dari media sosial. Sebanyak 200 komentar dijadikan sampel dan dianalisis menggunakan algoritma *K-Medoids* setelah melalui tahapan pra-pemrosesan data serta pembobotan dengan metode *Term Frequency–Inverse Document Frequency* (TF-IDF). Proses klusterisasi menghasilkan tiga kelompok utama sentimen, yaitu negatif, netral, dan positif.

Hasil analisis menunjukkan bahwa sentimen negatif mendominasi dengan total 122 komentar, sementara sentimen netral dan positif masing-masing terdiri dari 39 komentar. Temuan ini mengindikasikan bahwa sebagian besar pengguna media sosial menunjukkan reaksi negatif terhadap keputusan pemecatan STY. Respon tersebut terekam dalam bentuk kritik langsung, ketidaksetujuan terhadap kebijakan federasi, maupun dukungan terhadap figur Shin Tae-yong.

Visualisasi hasil pengujian yang diperoleh melalui implementasi Python disajikan sebagai berikut:

3.1. Dataset

Dataset yang digunakan merupakan data mentah berupa komentar warganet yang dikumpulkan dari berbagai sumber. Data ini akan digunakan dalam proses klasterisasi untuk membentuk tiga kelompok klaster. Tabel 1 menampilkan sebagian dari total 200 komentar warganet yang nantinya akan melalui proses pembersihan data sebelum dianalisis menggunakan algoritma klasterisasi.

Tabel 1 Dataset

No	Komentar
1	Kenapa malah dipecat? Dia sudah bawa perubahan besar loh.
2	Memang sudah saatnya, hasilnya kurang memuaskan.
3	Federasi aneh, STY itu pelatih terbaik yang pernah kita punya!
4	Terima kasih atas jasamu coach, semoga sukses di tempat lain.
5	Ah, drama lagi. Kapan sih sepak bola Indonesia serius?
6	Kalau lihat performa terakhir, sih wajar juga dia dipecat.
7	STY harusnya dikasih kesempatan lebih, jangan cuma lihat hasil akhir!
8	Kalau mau ganti, ganti juga sistem federasinya sekalian.
9	Pelatih top dunia pun susah kalau federasinya begini.
10	Mungkin ini keputusan terbaik untuk semua pihak.
...	
200	kita harus tetap dukung timnas, siapa pun pelatihnya.

3.2. Processing

Pada metode K-Medoids, istilah *processing* merujuk pada tahapan pengelompokan data ke dalam sejumlah klaster dengan memilih beberapa titik sebagai medoid atau pusat representatif klaster[6]. Selanjutnya, dilakukan perhitungan jarak antara tiap data dengan medoid yang ada, kemudian data dikelompokkan berdasarkan kedekatan jarak tersebut. Proses ini dilakukan secara berulang dengan mengganti medoid menggunakan titik lain dalam upaya mengurangi total biaya (cost) hingga diperoleh struktur klaster yang paling efisien.

a. *Cleansing*

Cleansing adalah tahap pertama dalam *preprocessing* teks yang bertujuan untuk membersihkan teks dari karakter-karakter yang tidak relevan[8]. B Rincian hasil preprocessing terhadap beberapa sampel komentar ditampilkan pada Tabel 2.

Tabel 2 Proses Cleaning

No	Sebelum	Sesudah
1	Kenapa malah dipecat? Dia sudah bawa perubahan besar loh.	Kenapa malah dipecat Dia sudah bawa perubahan besar loh
2	Memang sudah saatnya, hasilnya kurang memuaskan.	Memang sudah saatnya hasilnya kurang memuaskan

b. *Case Folding*

Case folding, yaitu mengubah semua huruf dalam teks menjadi huruf kecil. Proses case folding terhadap data ditampilkan pada Tabel 3.

Tabel 3 Proses Casefolding

No	Sebelum	Sesudah
----	---------	---------

1	Kenapa malah dipecat Dia sudah bawa perubahan besar loh	kenapa malah dipecat dia sudah bawa perubahan besar loh
2	Memang sudah saatnya hasilnya kurang memuaskan	memang sudah saatnya hasilnya kurang memuaskan

c. *Tokenizing*

Tokenizing yaitu memecah teks menjadi unit terkecil yang disebut token, biasanya berupa kata. Hasil tokenizing terhadap data ditampilkan pada Tabel 4.

Tabel 4 Proses Tokenizing

No	Sebelum	Sesudah
1	kenapa malah dipecat dia sudah bawa perubahan besar loh	['kenapa', 'malah', 'dipecat', 'dia', 'sudah', 'bawa', 'perubahan', 'besar', 'loh']
2	memang sudah saatnya hasilnya kurang memuaskan	['memang', 'sudah', 'saatnya', 'hasilnya', 'kurang', 'memuaskan']

d. *Stopword removal*

Stopword removal merupakan tahapan untuk menghapus kata-kata umum yang sering muncul tetapi tidak memiliki makna signifikan dalam analisis. Berikut dapat dilihat tokenizing data pada Tabel 5.

Tabel 5 Proses Stopword removal

No	Sebelum	Sesudah
1	['kenapa', 'malah', 'dipecat', 'dia', 'sudah', 'bawa', 'perubahan', 'besar', 'loh']	['kenapa', 'malah', 'dipecat', 'bawa', 'perubahan', 'besar']
2	['memang', 'sudah', 'saatnya', 'hasilnya', 'kurang', 'memuaskan']	['saatnya', 'hasilnya', 'kurang', 'memuaskan']

e. *Stemming*

stemming, yaitu mengubah kata-kata ke bentuk dasarnya (*stem/root word*). Berikut dapat dilihat hasil stemming data pada Tabel 6.

Tabel 6 Proses Stemming

No	Sebelum	Sesudah
1	['kenapa', 'malah', 'dipecat', 'bawa', 'perubahan', 'besar']	['kenapa', 'malah', 'pecat', 'bawa', 'ubah', 'besar']
2	['saatnya', 'hasilnya', 'kurang', 'memuaskan']	['saat', 'hasil', 'kurang', 'puas']

3.3. Pembobotan Dokumen TF-IDF

Term Weighting adalah teknik dalam pemrosesan teks yang digunakan untuk memberikan bobot atau nilai numerik pada setiap kata dalam suatu dokumen. Teknik ini bertujuan untuk menentukan seberapa penting suatu kata dalam suatu dokumen dibandingkan dengan kata-kata lainnya. Berikut dapat dilihat hasil pembobotan dokumen pada tabel 7.

Tabel 7 Pembobotan TF-IDF

Term	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
kenapa	1	0	0	0	0	0	0	0	0	0
malah	1	0	0	0	0	0	0	0	0	0

pecat	0,39794	0	0	0	0	0,39794	0	0	0	0
bawa	1	0	0	0	0	0	0	0	0	0
ubah	1	0	0	0	0	0	0	0	0	0
besar	1	0	0	0	0	0	0	0	0	0
saat	0	1	0	0	0	0	0	0	0	0
hasil	0	0,39794	0	0	0	0	0,39794	0	0	0
kurang	0	1	0	0	0	0	0	0	0	0
puas	0	1	0	0	0	0	0	0	0	0
federasi	0	0	0,045757	0	0	0,045757	0,045757	0	0	0
aneh	0	0	1	0	0	0	0	0	0	0
sty	0	0	0,39794	0	0	0	0,39794	0	0	0
timnas	0	0	0	0	0	0	0	0	0	1

3.4. K-Medoids

Pengelompokan dan Perhitungan Jarak merupakan tahap penting dalam proses *clustering* untuk mengelompokkan data berdasarkan kemiripan tertentu. Dalam metode *K-Medoids*.

a. Centroid random

Langkah awal dalam proses klasterisasi adalah pemilihan centroid secara acak, yaitu dengan menentukan titik pusat dari data tanpa mengikuti pola tertentu. Tahapan ini ditunjukkan pada Tabel 8.

Tabel 8 Centroid Random

Cluster	Medoid	kena pa	mala h	pecat	bawa	ubah	besar	saa t	hasil	kurang	pua s	federasi	ane h	sty	timnas
C1	D1	1	1	0,39794	1	1	1	0	0	0	0	0	0	0	0
C2	D3	0	0	0	0	0	0	0	0	0	0	0,045757	1	0,39794	0
C3	D10	0	0	0	0	0	0	0	0	0	0	0	0	0	1

b. Hitung jarak data ke *Euclidean Distance*

Metode K-Medoids mampu membentuk pengelompokan data yang konsisten dan cukup andal dalam menghadapi data pencilan (*outlier*). Hasil akhir proses klasterisasi berdasarkan perhitungan jarak antar data dan medoid disajikan pada Tabel 9.

Tabel 9 Euclidean Distance

Dokumen	Jarak C1	Jarak C2	Jarak C3	Min distance	Custer
D1	0	1,78	2,24	0	C1
D2	2,24	1,78	1,41	1,41	C3
D3	1,78	0	1	0	C2
D4	0	1,41	1	0	C1
D5	0	1,41	1	0	C1
D6	0,57	0,57	1	0,57	C2
D7	0,57	0,57	1	0,57	C2
D8	0	1,41	1	0	C1
D9	0	1,41	1	0	C1
D10	2,24	1,41	0	0	C3

c. Hasil Clustering

Hasil clustering merupakan hasil setelah semua proses. Dapat kita lihat pada table 10 dibawah ini:

Tabel 10 Hasil Clustering

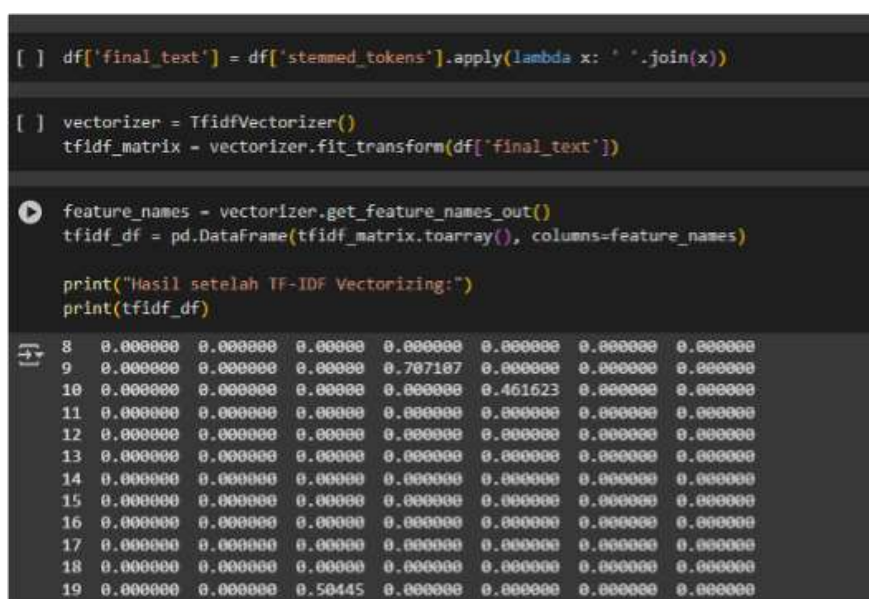
Cluster	Dokumen	Sentimen
C1	D1,D4,D5,D8,D8	Negatif
C2	D3,D6,D7	Netral
C3	D2,D10	Positif

4. Pengujian Pada Python

Pengujian dilakukan melalui sejumlah tahapan yang dirancang untuk memperoleh hasil klasterisasi secara optimal. Uraian dari masing-masing tahapan disajikan pada bagian berikut.

a. Pembobotan TF-IDF

Pemberian bobot kata (term weighting) adalah proses untuk menentukan seberapa penting suatu kata dalam dokumen berdasarkan frekuensi kemunculan atau relevansinya terhadap keseluruhan data. Proses ini dapat dilihat pada Gambar 2.



```
[ ] df['final_text'] = df['stemmed_tokens'].apply(lambda x: ' '.join(x))

[ ] vectorizer = TfidfVectorizer()
tfidf_matrix = vectorizer.fit_transform(df['final_text'])

feature_names = vectorizer.get_feature_names_out()
tfidf_df = pd.DataFrame(tfidf_matrix.toarray(), columns=feature_names)

print("Hasil setelah TF-IDF Vectorizing:")
print(tfidf_df)
```

	0	1	2	3	4	5	6	7
8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.707107	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000	0.461623	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
16	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
17	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
18	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
19	0.000000	0.000000	0.50445	0.000000	0.000000	0.000000	0.000000	0.000000

Gambar 2 Proses TF-IDF Python

b. Centroid Random

Inisialisasi centroid secara acak merupakan tahap awal dalam proses clustering, di mana sejumlah titik dipilih secara acak dari data sebagai pusat awal kelompok (centroid). Pemilihan ini memengaruhi hasil akhir pembentukan cluster dan akan diperbarui pada iterasi berikutnya hingga konvergen. Tahap ini ditunjukkan pada Gambar 3.



```

kmedoids = KMedoids(n_clusters=3, random_state=42)
kmedoids.fit(tfidf_matrix)

KMedoids
KMedoids(n_clusters=3, random_state=42)

[ ] df['Cluster'] = kmedoids.labels_

cluster_labels = {
    0: 'Negatif',
    1: 'Positif',
    2: 'Netral'
}
df['Cluster_Label'] = df['Cluster'].map(cluster_labels)

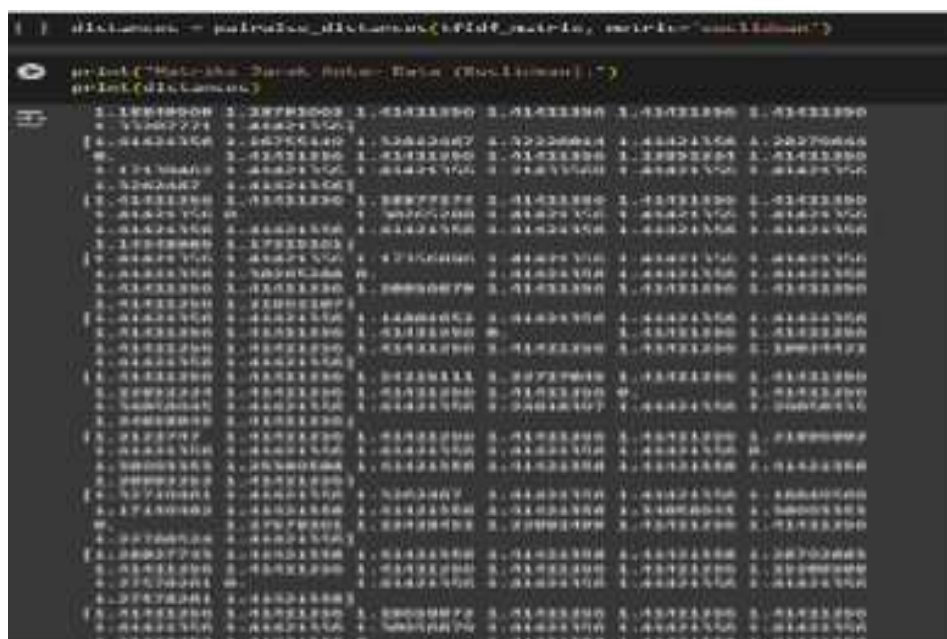
print("=== Hasil Clustering dengan Label ===")
print(df[['Komentar', 'Cluster', 'Cluster_Label']])
print("\n")

```

Gambar 3 Centroid Random Python

- c. Hitung jarak data ke *Euclidean Distance*

Perhitungan jarak Euclidean dilakukan dengan mengambil akar kuadrat dari jumlah selisih kuadrat antara dua titik data. Jarak yang lebih kecil menunjukkan kedekatan data dengan medoid, sehingga besar kemungkinan data tersebut tergabung dalam cluster yang sama. Proses ini ditampilkan pada Gambar 4.



```

distances = pairwise_distances(tfidf_matrix, metric='euclidean')

print("Matriks Jarak Antar Baris (Euclidean):")
print(distances)

```

Gambar 4 Hasil Euclidean Distance Python

- d. Hasil Cluster

Hasil clustering menunjukkan bahwa data telah dikelompokkan ke dalam beberapa cluster berdasarkan kemiripan jarak terhadap medoid. Setiap data tergabung dalam cluster dengan medoid terdekat, dan proses ini terus berlanjut hingga tidak terjadi perpindahan lagi. Hasil akhir dari proses pengelompokan ini ditampilkan pada Gambar 5.

```

for i in range(3):
    print(f'\nData pada Cluster {i}:')
    print(df[df['Cluster'] == i]['Komentar'])

Data pada Cluster 0:
0    Kenapa malah dipecat? Dia sudah bawa perubahan...
1    Memang sudah saatnya, hasilnya kurang memuaskan.
3    Terima kasih atas jasanya coach, semoga sukses ...
4    Ah, drama lagi. Kapan sih sepak bola Indonesia...
5    Kalau lihat performa terakhir, sih wajar juga ...
...
193   mungkin ada faktor politik juga di balik peme...
194   wajar dipecat, target-target banyak yang gagal.
195   salut untuk dedikasi shin tae-yong, sukses sel...
196   semoga cepat rebound dan timnas makin maju.
198   pecat shin tae-yong - mundur satu langkah.
Name: Komentar, length: 144, dtype: object

Data pada Cluster 1:
7    Kalau mau ganti, ganti juga sistem federasinya...
18   Pecat STY? Siapa penggantinya? Jangan lebih pa...
19   Pelatih bagus, tapi kalau terus gagal ya harus...
28   Saya setuju dia diganti, tapi harus yang lebih...
30   Siapapun pelatihnya, kalau sistem bobrok ya pe...
57   Kalau mau ganti, ganti juga sistem federasinya...
68   Pecat Shin Tae-yong? Siapa penggantinya? Janga...
69   Pelatih bagus, tapi kalau terus gagal ya harus...
78   Saya setuju dia diganti, tapi harus yang lebih...
80   Siapapun pelatihnya, kalau sistem bobrok ya pe...
107  Kalau mau ganti, ganti juga sistem federasinya...
118  Pecat Shin Tae-yong? Siapa penggantinya? Janga...
119  Pelatih bagus, tapi kalau terus gagal ya harus...
128  Saya setuju dia diganti, tapi harus yang lebih...
130  Siapapun pelatihnya, kalau sistem bobrok ya pe...
157  Kalau mau ganti, ganti juga sistem federasinya...
168  Pecat shin tae-yong? siapa penggantinya? janga...
169  pelatih bagus, tapi kalau terus gagal ya harus...

```

Gambar 5 Hasil Clustering

4. KESIMPULAN

Studi ini berhasil mengidentifikasi pola sentimen masyarakat terhadap pemberhentian Shin Tae-yong dengan memanfaatkan algoritma K-Medoids. Temuan menunjukkan bahwa mayoritas komentar di media sosial tergolong dalam kategori sentimen negatif, yaitu sekitar 61%. Penggunaan K-Medoids terbukti mampu mengelompokkan data teks tak terstruktur secara cukup optimal, yang ditunjukkan oleh nilai *Silhouette Score* sebesar 0,41.

REFERENSI

- [1] L. M. Barus, H. G. Simanullang, and A. P. Silalahi, "Classification of English-Language Racial Hate Speech Using Random Forest", *J. Intell. Comput. Inf. Syst.*, vol. 1, no. 1, pp. 1–17, 2025.
- [2] N. Duha, H. G. Simanullang, and A. P. Silalahi, "ANALISIS PENGARUH VARIASI NILAI P PADA METODE MINKOWSKI DISTANCE DALAM MENENTUKAN KEMIRIPAN ABSTRAK SKRIPSI Harlen Gilbert Simanullang, Arina Prima Silalahi □, Nadyarni Natalis Caesarin Duha," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 9, no. 2, pp. 255–263, 2025.
- [3] M. I. Hutapea and A. P. Silalahi, "Moderna's Vaccine Using the K-Nearest Neighbor (KNN) Method: An Analysis of Community Sentiment on Twitter," *J. Penelit. Pendidik. IPA*, vol. 9, no. 5, pp. 3808–3814, 2023, doi: 10.29303/jppipa.v9i5.3203.
- [4] H. G. Simanullang, S. Wijono, B. Kristianto, W. Sulistyono, and A. P. Silalahi, "Detecting Cyberbullying Using Machine Learning Approaches," in *2024 Ninth International Conference on Informatics and Computing (ICIC)*, 2024, pp. 1–8. doi: 10.1109/ICIC64337.2024.10956842.
- [5] H. D. Tampubolon, S. Suhada, M. Safii, S. Solikhun, and D. Suhendro, "Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi," *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 6–12, 2021, doi: 10.35960/ikomti.v2i2.703.

-
- [6] H. Utami, "Analisis Sentimen dari Aplikasi Shopee Indonesia Menggunakan Metode Recurrent Neural Network," *Indones. J. Appl. Stat.*, vol. 5, no. 1, p. 31, 2022, doi: 10.13057/ijas.v5i1.56825.