



Klasterisasi Komentar Bermuatan Pelanggaran Etika Digital terhadap Menteri ESDM pada Media Sosial X Menggunakan Metode K-Means

Amalia Azzahra¹, Dhaifina Nabila², Lusiyanti³, Fazril Ramadhan⁴, Abdul Syukur⁵, Eko Sutono⁶

¹⁻⁶ Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: ¹aamalyyazzra@gmail.com, ²dhaifinanab@gmail.com, ³lusiy2610@gmail.com,
⁴fazrilrmd8@gmail.com, ⁵aabdulsyukur77@gmail.com, ⁶dosen02598@unpam.ac.id

Abstrak—Media sosial X menjadi sarana masyarakat dalam menyampaikan opini terhadap tokoh publik, namun juga memunculkan komentar bermuatan pelanggaran etika digital seperti hinaan dan ujaran kebencian. Penelitian ini bertujuan untuk mengelompokkan komentar netizen terhadap Menteri ESDM menggunakan metode K-Means. Data diperoleh melalui proses scraping sebanyak 500 tweet. Tahapan penelitian meliputi preprocessing, pembobotan TF-IDF, dan clustering menggunakan K-Means. Hasil penelitian menunjukkan bahwa komentar berhasil dikelompokkan menjadi tiga cluster, yaitu komentar netral/kritik biasa, komentar hinaan atau kata kasar, dan komentar provokasi/ujaran kebencian. Metode K-Means mampu membantu analisis komentar bermuatan pelanggaran etika digital pada media sosial.

Kata Kunci: K-Means, Text Mining, Media Sosial X, Etika Digital, Clustering.

Abstract—*Social media X is widely used by the public to express opinions toward public figures, but it also contains comments that violate digital ethics such as insults and hate speech. This study aims to classify comments directed at the Minister of Energy and Mineral Resources using the K-Means method. The data were collected through scraping techniques with a total of 500 tweets. The research stages included preprocessing, TF-IDF weighting, and clustering using K-Means. The results showed that the comments were grouped into three clusters: neutral comments, insulting comments, and hate speech comments. The K-Means method was able to assist in analyzing comments containing violations of digital ethics on social media.*

Keywords: *K-Means, Text Mining, Social Media X, Digital Ethics, Clustering.*

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi menyebabkan media sosial menjadi salah satu sarana utama masyarakat dalam menyampaikan opini dan berinteraksi secara digital. Media sosial tidak hanya digunakan sebagai media berbagi informasi, tetapi juga menjadi ruang publik untuk menyampaikan kritik, pendapat, dan ekspresi terhadap isu sosial maupun tokoh publik. Salah satu platform yang paling aktif digunakan dalam percakapan publik adalah media sosial X atau Twitter karena memiliki karakteristik penyebaran informasi yang cepat dan real-time (Ramdhani et al., 2025).

Tingginya aktivitas komunikasi pada media sosial juga memunculkan dampak negatif berupa komentar bermuatan ujaran kebencian, hinaan, provokasi, dan pelanggaran etika digital. Pengguna media sosial sering kali menyampaikan pendapat tanpa mempertimbangkan etika komunikasi digital sehingga memunculkan hate speech terhadap individu maupun kelompok tertentu. Hate speech merupakan bentuk komunikasi yang mengandung penghinaan, diskriminasi, maupun serangan verbal yang dapat menimbulkan konflik sosial (Refa et al., 2025).

Menteri Energi dan Sumber Daya Mineral (ESDM) sebagai tokoh publik sering menjadi perhatian masyarakat di media sosial X. Berbagai komentar yang muncul tidak hanya berupa kritik terhadap kebijakan, tetapi juga mengandung kata kasar, hinaan, serta ujaran kebencian. Fenomena tersebut menunjukkan bahwa media sosial dapat menjadi sarana penyebaran pelanggaran etika digital apabila digunakan tanpa kontrol dan tanggung jawab dalam berkomunikasi (Refa et al., 2025).

Dalam bidang data mining, analisis komentar media sosial dapat dilakukan menggunakan teknik text mining untuk mengidentifikasi pola komentar berdasarkan kemiripan kata. Salah satu metode yang sering digunakan adalah K-Means Clustering. Metode K-Means merupakan algoritma unsupervised learning yang digunakan untuk mengelompokkan data ke dalam beberapa cluster



berdasarkan karakteristik tertentu. Metode ini dinilai efektif dalam mengelompokkan data teks dalam jumlah besar dengan proses komputasi yang relatif cepat dan efisien (Larasati et al., 2025).

Penelitian sebelumnya menunjukkan bahwa metode K-Means mampu digunakan dalam analisis sentimen dan pengelompokan komentar negatif pada media sosial Twitter. Penelitian mengenai komentar hate speech menggunakan algoritma K-Means menunjukkan bahwa proses clustering mampu mengelompokkan komentar ke dalam kategori negatif, netral, dan positif berdasarkan kemiripan pola kata dan konteks komentar (Ramdhani et al., 2025)

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk melakukan klusterisasi komentar bermuatan pelanggaran etika digital terhadap Menteri ESDM pada media sosial X menggunakan metode K-Means. Penelitian ini diharapkan dapat membantu memahami pola komentar netizen serta memberikan gambaran mengenai bentuk pelanggaran etika digital yang terjadi pada media sosial.

2. METODE

2.1 Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis untuk memperoleh hasil clustering komentar bermuatan pelanggaran etika digital pada media sosial X. Proses penelitian dimulai dari pengumpulan data tweet menggunakan teknik scraping pada platform X dengan kata kunci yang berkaitan dengan Menteri ESDM. Setelah data berhasil diperoleh, dilakukan tahap preprocessing untuk membersihkan data teks agar siap digunakan dalam proses analisis.

Tahap preprocessing meliputi cleansing, case folding, tokenizing, stopword removal, dan stemming. Selanjutnya data teks diubah menjadi data numerik menggunakan metode TF-IDF (Term Frequency–Inverse Document Frequency). Data yang telah diproses kemudian dianalisis menggunakan algoritma K-Means Clustering untuk mengelompokkan komentar berdasarkan kemiripan pola kata. Hasil clustering dibagi menjadi tiga cluster yaitu komentar netral/kritik biasa, komentar mengandung hinaan atau kata kasar, dan komentar provokasi atau ujaran kebencian. Tahap terakhir dilakukan analisis hasil cluster dan penarikan kesimpulan penelitian (Riadi et al., 2023).

2.2 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini berupa data komentar atau tweet pada media sosial X yang berkaitan dengan Menteri ESDM. Data diperoleh melalui proses scraping menggunakan tools tweet-harvest dengan kata kunci tertentu yang relevan dengan topik penelitian. Jumlah dataset yang digunakan sebanyak 500 tweet berbahasa Indonesia (Difyah et al., 2025).

Dataset penelitian terdiri dari beberapa atribut, antara lain:

- a. *username*
- b. *full_text*
- c. *reply_count*
- d. *retweet_count*
- e. *favorite_count*
- f. *created_at*

Dari seluruh atribut tersebut, penelitian ini berfokus pada atribut *full_text* karena digunakan sebagai data utama dalam proses text mining dan clustering. Data tweet yang diperoleh kemudian melalui tahapan preprocessing sebelum dianalisis menggunakan metode K-Means.

2.3 Preprocessing Data

Preprocessing data merupakan tahap penting dalam text mining yang bertujuan untuk membersihkan dan menyiapkan data teks agar dapat diproses dengan baik oleh sistem. Tahapan preprocessing yang digunakan dalam penelitian ini meliputi (Mauliza & Sipayung, 2024):

- a. *Case Folding*: Tahap awal di mana teks ulasan diubah menjadi huruf kecil (diseragamkan) dengan mengubah karakter huruf kapital menjadi huruf kecil.
- b. *Cleansing* (Pembersihan): Proses menghapus komponen yang dianggap sebagai gangguan (*noise*), seperti URL, *@mention*, *#hashtag*, tautan (*link*), spasi ganda, emotikon, serta angka.



- c. *Tokenizing*: Proses memotong-motong kalimat atau teks menjadi kata per kata yang disebut sebagai token.
- d. *Filtering (Stopword Removal)*: Proses menghapus kata-kata yang dianggap tidak penting atau tidak memberikan pengaruh besar terhadap proses klasifikasi, seperti kata hubung ("dan", "atau", "ke").
- e. *Stemming*: Tahap untuk mengubah kata-kata berimbuhan kembali ke dalam bentuk dasarnya menggunakan kamus kata dasar.

2.4 Transformasi Data

Setelah preprocessing selesai dilakukan, data teks diubah ke dalam bentuk numerik menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Metode ini merupakan salah satu metode pembobotan kata (*word weighting*) yang digunakan untuk menghitung seberapa penting nilai sebuah kata di dalam suatu dokumen. Proses komputasi pada metode ini dilakukan dengan cara mengalikan frekuensi kemunculan suatu kata (*Term Frequency*) dengan inversi frekuensi dokumen tersebut (*Inverse Document Frequency*) (Siddiq et al., 2025).

2.5 K-Means Clustering

Metode K-Means Clustering merupakan salah satu algoritma data mining yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik data. Algoritma ini bekerja dengan cara menentukan sejumlah cluster kemudian memilih centroid sebagai pusat cluster. Setiap data akan dihitung jaraknya terhadap centroid dan ditempatkan pada cluster dengan jarak terdekat. Proses tersebut dilakukan secara berulang hingga centroid stabil dan tidak mengalami perubahan lagi (Ramadhan & Namung, 2023).

2.6 Euclidean Distance

Euclidean Distance merupakan metode perhitungan jarak yang digunakan untuk mengukur jarak antara dua titik dalam ruang multidimensi. Pada metode K-Means, Euclidean Distance digunakan untuk menghitung kedekatan data terhadap centroid cluster sehingga data dapat dikelompokkan ke cluster terdekat (Priadi et al., 2022).

3. ANALISA DAN PEMBAHASAN

Dalam penelitian ini, metode klasterisasi K-Means digunakan untuk mengklasifikasikan komentar yang bermuatan pada pelanggaran etika digital terhadap Menteri ESDM di media sosial X. Dataset yang digunakan terdiri dari 500 tweet yang dikumpulkan melalui scraping kata kunci yang terkait dengan kata kunci "Bahlil" dan "Menteri ESDM".

Tahapan penelitian meliputi preprocessing data, pembobotan kata menggunakan TF-IDF, dan proses clustering menggunakan algoritma K-Means. Hasil clustering menghasilkan tiga kelompok cluster yaitu C1, C2, dan C3 berdasarkan kemiripan pola kata pada komentar.

3.1 Hasil Preprocessing Data

Tahap preprocessing dilakukan untuk membersihkan data teks agar dapat diproses dengan baik pada tahap clustering. Tahapan preprocessing terdiri dari cleansing, case folding, tokenizing, stopword removal, dan stemming.

3.1.1 Cleansing

Tahap cleansing dilakukan untuk menghapus karakter yang tidak diperlukan seperti URL, emoji, mention, hashtag, angka, dan tanda baca.

Tabel 1. Tahap Cleansing

Data Asli	Hasil Cleansing
“@mintchouchou marah gak si? doyoung dikata mirip bahlil”	“marah gak si doyoung dikata mirip bahlil”
“ https://t.co/cVn359HUzq Aowkaokw bakar sate sambil bahas MBG mas bahlil ganteng”	“Aowkaokw bakar sate sambil bahas MBG mas bahlil ganteng”



“GOLKAR : Ramai Lagu Mas Bahlil Ganteng
di Medsos <https://t.co/Fs03ilqPaf>”

“GOLKAR Ramai Lagu Mas Bahlil
Ganteng di Medsos”

Berdasarkan hasil cleansing, data menjadi lebih bersih karena karakter yang tidak diperlukan telah dihapus.

3.1.2 Case Folding

Tahap case folding dilakukan dengan mengubah seluruh huruf menjadi huruf kecil (*lowercase*).

Tabel 2. Tahap case folding

Sebelum	Sesudah
“MBG Mas Bahlil Ganteng”	“mbg mas bahlil ganteng”
“STOP. PAKE. LAGU. BAHLIL.”	“stop pake lagu bahlil”
“Mas Bahlil Gendeng”	“mas bahlil gendeng”

3.1.3 Tokenizing

Tahap tokenizing dilakukan untuk memecah kalimat menjadi token kata.

Tabel 3. Tahap tokenizing

Kalimat	Hasil Tokenizing
“mbg mas bahlil ganteng”	[mbg, mas, bahlil, ganteng]
“stop pake lagu bahlil”	[stop, pake, lagu, bahlil]”
”maksudnya satir malah dianggap memuji”	[maksudnya, satir, malah, dianggap, memuji]

3.1.4 Stopword Removal

Tahap stopword removal dilakukan untuk menghapus kata umum yang tidak memiliki makna penting.

Tabel 4. Tahap stopword removal

Sebelum	Sesudah
“maksudnya satir malah dianggap memuji”	“satir dianggap memuji”
“rekomendasi film horror tapi gamau yg ada bahlil nya”	“rekomendasi film horror gamau bahlil”
“lagunya terus muncul di media sosial”	“lagu muncul media sosial”

3.1.5 Stemming

Tahap stemming dilakukan untuk mengubah kata berimbuhan menjadi kata dasar.

Tabel 5. Tahap stemming

Sebelum	Sesudah
“dianggap memuji”	“anggap puji”
“bersiliweran orang”	“siliwer orang”
“menikmati hari raya”	“nikmat hari raya”

Berdasarkan hasil preprocessing, data komentar menjadi lebih bersih dan seragam sehingga meningkatkan kualitas proses analisis pada tahap clustering.

3.2 Pembobotan Kata Menggunakan TF-IDF

Setelah preprocessing selesai, dilakukan pembobotan kata menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*). Metode ini digunakan untuk mengetahui tingkat kepentingan suatu kata dalam dokumen komentar.

3.2.1 Perhitungan TF-IDF

Berdasarkan hasil perhitungan pada dataset diperoleh:

- Jumlah seluruh dokumen (N) = 502
- Jumlah dokumen yang mengandung term “satir” (df) = 8
- Frekuensi term “satir” pada dokumen (TF) = 1

Maka perhitungan IDF untuk term “satir” yaitu:

$$IDF(\text{satir}) = \text{Log}_{10} \frac{502}{8} = 1.79$$

Selanjutnya dilakukan perhitungan TF-IDF:

$$TF-IDF(\text{satir}) = 1 \times 1.79 = 1.79$$

Selain term “satir”, dilakukan juga perhitungan pada beberapa term lain yang terdapat pada dataset penelitian.

Tabel 6. Hasil perhitungan TF-IDF

Term	TF	df	IDF	TF-IDF
bahlil	1	497	0.004	0.004
mbg	1	123	0.610	0.610
ganteng	1	152	0.518	0.518
satir	1	8	1.798	1.798
goblok	1	9	1.746	1.746

Perhitungan IDF untuk term “bahlil” diperoleh sebagai berikut:

$$IDF(\text{bahlil}) = \text{Log}_{10} \frac{502}{497} = 0.004$$

Nilai IDF term “bahlil” lebih kecil dibandingkan term lain karena kata tersebut muncul hampir pada seluruh dokumen sehingga dianggap kurang spesifik dalam merepresentasikan isi komentar.

Sedangkan term seperti “satir” dan “goblok” memiliki nilai IDF lebih tinggi karena hanya muncul pada sebagian kecil dokumen. Hal tersebut menunjukkan bahwa term tersebut lebih spesifik dan memiliki pengaruh lebih besar dalam proses clustering menggunakan metode K-Means.

3.3 Proses Clustering Menggunakan K-Means

Setelah dilakukan pembobotan kata menggunakan metode TF-IDF, tahap selanjutnya yaitu proses clustering menggunakan algoritma K-Means. Metode K-Means digunakan untuk mengelompokkan komentar berdasarkan kemiripan pola kata pada data tweet.

Pada penelitian ini digunakan jumlah cluster sebanyak tiga cluster yaitu C1, C2, dan C3. Penentuan jumlah cluster dilakukan berdasarkan karakteristik komentar yang ditemukan pada dataset, yaitu komentar umum, komentar satire/kritik, dan komentar dengan bahasa agresif.

3.3.1 Penentuan Jumlah Cluster

Pada penelitian ini jumlah cluster ditentukan sebanyak tiga cluster (k = 3). Penentuan jumlah cluster dilakukan berdasarkan karakteristik komentar yang ditemukan pada dataset penelitian, yaitu:

Tabel 7. Penentuan jumlah cluster data

Cluster	Karakteristik
C1	Komentar umum dan candaan ringan
C2	Komentar satire dan kritik
C3	Komentar agresif dan pelanggaran etika digital



Penentuan jumlah cluster sebanyak tiga cluster bertujuan agar proses pengelompokan komentar dapat menghasilkan kategori yang lebih jelas berdasarkan pola kata dan konteks komentar pada media sosial X.

3.3.2 Penentuan Centroid Awal

Setelah menentukan jumlah cluster, tahap berikutnya yaitu menentukan centroid awal. Centroid merupakan titik pusat cluster yang digunakan sebagai acuan dalam proses pengelompokan data. Pada penelitian ini centroid awal dipilih secara acak dari data hasil pembobotan TF-IDF.

Tabel 8. Centroid awal cluster

Cluster	Term satir	Term goblok
C1	1.798	0.000
C2	0.000	1.746
C3	0.610	0.518

Nilai centroid awal tersebut digunakan untuk menghitung jarak setiap data tweet terhadap masing-masing cluster menggunakan metode Euclidean Distance.

3.3.3 Perhitungan Euclidean Distance

Perhitungan jarak antar data dilakukan menggunakan Euclidean Distance dengan rumus:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

$d(x,y)$ = jarak antar data

x_i = nilai data ke- i

y_i = nilai centroid ke- i

n = jumlah atribut

Sebagai contoh digunakan bobot TF-IDF dari term “satir” dan “goblok” yang terdapat pada dataset penelitian.

Tabel 9. Nilai TF-IDF pada data tweet

Term	Nilai TF-IDF
satir	1.798
goblok	1.746

Sedangkan centroid cluster C2 memiliki nilai:

Tabel 10. Nilai centroid awal cluster c2

Term	Nilai Centroid
satir	0.000
goblok	1.746

Maka perhitungan Euclidean Distance dilakukan sebagai berikut.

- Menghitung Selisih Nilai Data dan Centroid
Untuk term “satir”: $1.798 - 0.000 = 1.798$
Untuk term “goblok”: $1.746 - 1.746 = 0$
- Menghitung Kuadrat Selisih
Untuk term “satir”: $1.798^2 = 3.232$
Untuk term “goblok”: $0^2 = 0$
- Menjumlahkan Hasil Kuadrat
 $3.232 + 0 = 3.232$



d. Menghitung Akar Kuadrat

$$\sqrt{3.232} = 1.798$$

Berdasarkan hasil perhitungan diperoleh nilai Euclidean Distance sebesar 1.798. Nilai jarak tersebut digunakan untuk menentukan posisi data tweet pada cluster dengan centroid terdekat.

3.3.4 Hasil Iterasi K-Means

Setelah dilakukan perhitungan jarak menggunakan Euclidean Distance, setiap data tweet dikelompokkan ke cluster dengan jarak terdekat. Proses clustering dilakukan secara berulang hingga tidak terjadi perubahan posisi data pada cluster.

Pada setiap iterasi, centroid baru dihitung berdasarkan rata-rata nilai TF-IDF dari seluruh data pada masing-masing cluster. Proses iterasi berhenti ketika hasil cluster telah stabil.

Pada penelitian ini jumlah cluster ditentukan sebanyak tiga cluster ($k = 3$). Namun berdasarkan hasil clustering pada dataset penelitian, sebagian besar data terkelompok pada cluster C1 dan C2, sedangkan cluster C3 hanya memiliki sangat sedikit anggota data.

Hasil clustering menunjukkan bahwa komentar pada media sosial X lebih banyak didominasi komentar umum, candaan, satire, dan kritik dibandingkan komentar agresif.

3.4 Analisis Hasil Clustering

3.4.1 Cluster C1

Cluster C1 didominasi komentar umum dan candaan ringan terkait fenomena viral pada media sosial X.

Tabel 11. Hasil cluster c1

No.	Komentar
1.	“anak buah kanda bahlil”
2.	“hidup bahlil”
3.	“mbg mas bahlil ganteng buah apa yang paling manis buahlil”

Komentar pada cluster ini cenderung berupa candaan, opini umum, dan pembahasan ringan tanpa unsur agresif secara langsung.

3.4.2 Cluster C2

Cluster C2 berisi komentar satire dan kritik terhadap fenomena yang sedang dibahas pada media sosial X.

Tabel 12. Hasil cluster c2

No.	Komentar
1.	“maksudnya satir malah dianggap memuji”
2.	“lagunya terus muncul di media sosial”
3.	“stop jadiin pejabat inkompeten as a joke”

Komentar pada cluster ini menunjukkan bentuk satire dan kritik sosial yang disampaikan netizen pada media sosial.

3.4.3 Cluster C3

Cluster C3 hanya memiliki sedikit data sehingga tidak menunjukkan pola komentar yang dominan. Oleh karena itu, analisis pada cluster ini tidak dilakukan secara mendalam.

4. KESIMPULAN

Berdasarkan hasil penelitian, metode K-Means Clustering berhasil mengelompokkan 500 komentar netizen terkait Menteri ESDM di media sosial X ke dalam tiga cluster, yaitu komentar umum dan candaan ringan (C1), komentar satire dan kritik (C2), serta komentar agresif yang mengandung pelanggaran etika digital (C3). Proses preprocessing dan pembobotan TF-IDF membantu



JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 4, No. 3 Tahun 2026
ISSN 3025-0919 (media online)
Hal 763-770

meningkatkan kualitas data sehingga clustering dapat dilakukan dengan baik. Hasil penelitian menunjukkan bahwa sebagian besar komentar didominasi oleh candaan, opini umum, dan kritik dibandingkan ujaran kebencian. Dengan demikian, metode K-Means efektif digunakan untuk menganalisis pola komentar serta mengidentifikasi bentuk pelanggaran etika digital pada media sosial.

REFERENCES

- Difyah, S. M. Z. A., Winanto, E. A., & Jusia, P. A. (2025). Analisis Sentimen Persepsi Publik terhadap Tagar #KaburAjaDulu di Twitter Menggunakan Metode Lexicon-Based dengan Kamus InSet. *Jurnal Processor*, 20(2), 223–233.
- Mauliza, R. N., & Sipayung, Y. R. (2024). Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes. *Technomedia Journal*, 9(1), 1–16. <https://doi.org/10.33050/tmj.v9i1.2212>
- Larasati, A. Z. M., Winarsih, N. A. S., Rohman, M. S., & Saraswati, M. S. (2025). Penerapan Metode K-Means Clustering dalam Menganalisis Sentimen Masyarakat terhadap K-Popers pada Twitter. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 201–210. <https://databoks.katadata.co.id/>
- Pribadi, W. W., Yunus, A., & Wiguna, A. S. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 493–500. <https://doi.org/10.36040/jati.v6i2.4808>
- Ramadhan, V. P., & Namung, G. M. (2023). Klasterisasi Komentar Cyberbullying Masyarakat di Instagram berdasarkan K-Means Clustering. *J-Intech*, 11(1), 32–39. <https://doi.org/10.32664/j-intech.v11i1.846>
- Ramdhani, M. R. T., Budika, M. R. P., Santoso, M. V., Alfira, N., & Zahry, N. (2025). Analisis Sentimen Terhadap Komentar Negatif (Hate Speech) Di Twitter Dengan Algoritma K-Means Clustering Menggunakan Rapidminer. *Journal of Information Technology and Informatics Engineering*, 1(1), 57–61. <https://journal.jci.co.id/jitie/id/article/view/139/24>
- Riadi, I., Fadlil, A., & Murni, M. (2023). Identifying Hate Speech in Tweets with Sentiment Analysis on Indonesian Twitter Utilizing Support Vector Machine Algorithm. *Khazanah Informatika : Jurnal Ilmu Komputer Dan Informatika*, 9(2), 179–191. <https://doi.org/10.23917/khif.v9i2.22470>
- Siddiq, M. J., Jayasri, S., Suhendi, A., Hidayat, T., & Rizky, R. (2025). Analisis Sentimen Opini Masyarakat Terhadap Pilkada 2024 Di Media Sosial Twitter Menggunakan Algoritma Naive Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2), 609–622. <https://doi.org/10.23960/jitet.v13i2.6280>