

PENERAPAN METODE DECISION TREE PADA SENTIMEN MEDIA SOSIAL TERKAIT KOMISI PEMILIHAN UMUM (KPU) SEBELUM DAN SESUDAH PILPRES 2024

Yolanda Veren*¹, Sunneng Sandino Berutu², Haeni Budiati³

1. Universitas Kristen Immanuel, Indonesia
2. Universitas Kristen Immanuel, Indonesia
3. Universitas Kristen Immanuel, Indonesia

Article Info

Kata Kunci: Decision Tree; Komisi Pemilihan Umum (KPU); Sentimen publik; Pilpres 2024

Keywords: *Decision Tree; General Election Commission (KPU); Public sentiment; Presidential Election 2024*

Article history:

Received 29 September 2024

Revised 13 Oktober 2024

Accepted 4 November 2024

Available online 4 December 2024

DOI :

<https://doi.org/10.29100/jipi.v9i4.5532>

* Corresponding author.

Yolanda Veren

E-mail address:

yveren0152@gmail.com

ABSTRAK

Pemilihan Presiden 2024 di Indonesia menandai momen krusial dalam dinamika politik negara. Penelitian ini mengamati perubahan pandangan masyarakat terhadap Komisi Pemilihan Umum (KPU) sebelum dan setelah Pilpres 2024 dengan menerapkan metode klasifikasi *Decision Tree*. Dataset yang digunakan terdiri dari 5794 (tambahin tanggal) tweet yang telah melewati serangkaian tahap, mulai dari crawling, pre processing data, terjemahan, pelabelan, klasifikasi, hingga evaluasi model. Hasil analisis menunjukkan bahwa sebelum Pilpres 2024, sentimen positif mencapai 39.98%, sentimen netral sebesar 38.39%, dan sentimen negatif sebesar 21.63%. Setelah Pilpres 2024, terjadi peningkatan dalam sentimen positif menjadi 45.45%, sedangkan sentimen netral mengalami penurunan menjadi 36.67%, dan sentimen negatif juga mengalami penurunan menjadi 17.87%. Berdasarkan dataset Komisi Pemilihan Umum (KPU) setelah Pilpres 2024 akurasi prediksi meningkat menjadi 67, dengan sentimen netral memiliki presisi sebesar 70. Sedangkan berdasarkan dataset pada Komisi Pemilihan Umum (KPU) setelah Pilpres 2024, recall untuk sentimen positif adalah 69, sementara f1-score untuk sentimen netral adalah 72.

ABSTRACT

The Presidential Election of 2024 in Indonesia marks a crucial moment in the country's political dynamics. This study observes the changes in public perception towards the General Election Commission (KPU) before and after the 2024 Presidential Election by applying the Decision Tree classification method. The dataset used consists of 5798 tweets that have undergone a series of stages, ranging from crawling, pre-processing, translation, labeling, classification, to model evaluation. The analysis results indicate that before the 2024 Presidential Election, positive sentiment reached 39.98%, neutral sentiment was at 38.39%, and negative sentiment was at 21.63%. After the 2024 Presidential Election, there was an increase in positive sentiment to 45.45%, while neutral sentiment decreased to 36.67%, and negative sentiment also decreased to 17.87%. Prediction accuracy increased to 67, with neutral sentiment having a precision of 70. In the dataset focused on the KPU after the 2024 Presidential Election, the recall for positive sentiment was 69, while the f1-score for neutral sentiment was 72, using the KPU dataset before the 2024 Presidential Election.

I. PENDAHULUAN

DALAM rangka penyelenggaraan pemilihan presiden tahun 2024, KPU menegaskan komitmennya untuk mematuhi ketentuan hukum yang berlaku, terutama Pasal 167 ayat (1) UU Nomor 7 Tahun 2017 dan Pasal 201 ayat (8) UU Nomor 10 Tahun 2016, yang menetapkan jadwal pelaksanaan Pemilu dan Pemilihan serentak nasional. Dalam sistem demokrasi, semua warga memiliki hak yang sama dalam membuat keputusan penting dalam kehidupan kewarganegaraan. Melalui demokrasi, masyarakat Indonesia dapat berpartisipasi dalam pembentukan, penyusunan hukum, dan perkembangan, baik melalui perantara perwakilan maupun partisipasi langsung. Indonesia, sebagai negara yang menganut prinsip demokrasi, menunjukkan komitmen ini melalui

pelaksanaan pemilihan umum secara berkala [1].

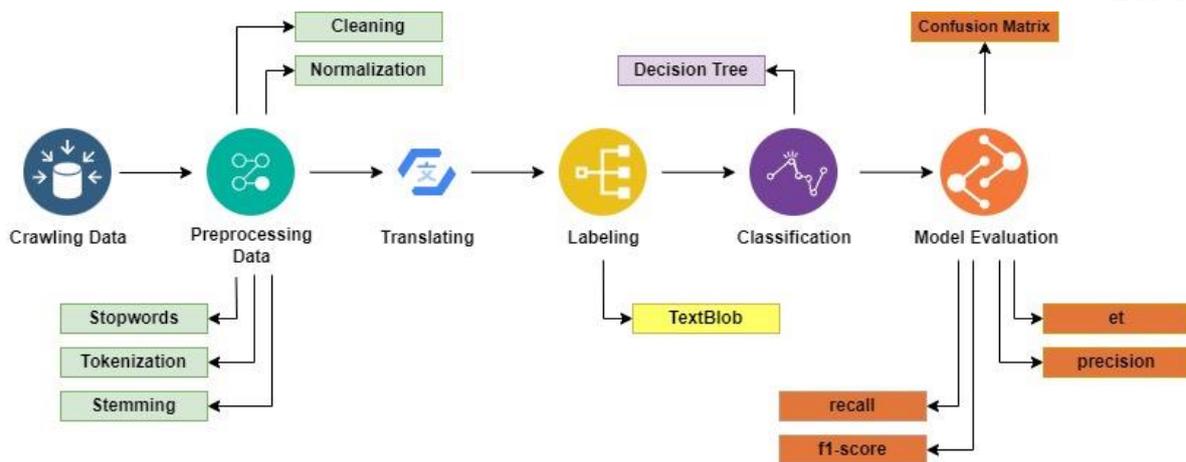
Perubahan besar dalam teknologi internet telah mempengaruhi cara individu atau kelompok berkomunikasi dalam berbagai bidang kehidupan. Kemunculan internet telah memicu pertumbuhan platform digital, termasuk media sosial, yang memungkinkan pengguna untuk berinteraksi, berkolaborasi dengan orang lain, menyajikan diri dan berinteraksi secara digital [2]. Selain itu, media sosial di Indonesia juga dimanfaatkan sebagai sarana untuk memberikan masukan atau kritik terhadap layanan yang disediakan oleh pemerintah. Terutama di platform Twitter, banyak orang yang menggunakan platform ini untuk memberikan pujian, kritik, atau bahkan tidak mengungkapkan pendapat terhadap kebijakan yang diambil oleh pemerintah saat ini [3]. Walaupun perkembangan jaringan sosial media memungkinkan masyarakat untuk berekspresi dan berpendapat secara bebas, masyarakat harus tetap mematuhi batasan yang ditetapkan dalam Pasal 27 ayat 3 UU ITE mengatur tentang larangan penyalahgunaan dan penyebaran informasi yang dapat merugikan orang lain [4].

Sebelumnya, telah dilakukan sejumlah penelitian terkait analisis sentimen, termasuk di antaranya adalah penelitian mengenai klasifikasi sentimen pengguna Twitter pada Pemilihan Presiden Indonesia tahun 2019 menggunakan metode Naive Bayes. Ada 240 data latih yang terdiri dari 134 sentimen negatif dan 106 sentimen positif. Penelitian ini menemukan bahwa penggunaan algoritma klasifikasi Naive Bayes pada data tweet menghasilkan tingkat akurasi sebesar 73%, dengan presisi untuk kelas negatif mencapai 78% dan presisi untuk kelas positif sebesar 66% [5].

Media sosial kini berperan signifikan dalam membentuk opini publik tentang isu politik, termasuk pandangan terhadap KPU dan pemilihan umum. Penelitian ini penting karena dapat mengisi celah pengetahuan dalam literatur analisis sentimen politik, memberikan wawasan praktis bagi pemangku kepentingan politik dan kampanye, serta relevan dengan isu-isu kontemporer terkait penggunaan teknologi dalam pemilu dan opini publik di media sosial. Dengan menganalisis perubahan opini dan tren publik seputar KPU melalui decision tree, penelitian ini dapat memberikan pandangan yang lebih mendalam tentang dinamika sosial politik yang sedang berlangsung [6]. Suatu pendekatan yang dapat diterapkan dalam analisis sentimen yaitu Decision Tree, yang merupakan salah satu teknik dari bidang pembelajaran mesin (machine learning). Decision Tree (pohon keputusan) adalah metode dalam data mining dan machine learning yang menghasilkan model prediksi menggunakan struktur pohon, di mana setiap node internal merepresentasikan tes pada atribut, cabang-cabangnya mewakili hasil dari tes tersebut, dan daun node mewakili nilai target atau kelas [7]. Tujuan dari penelitian ini adalah untuk menyelidiki sentimen yang tercermin dalam tweet di Twitter, baik yang bersifat positif maupun negatif, sebagai indikator kepercayaan masyarakat terhadap KPU sebagai penyelenggara pemilu tahun 2024 [8]. Penelitian ini akan fokus pada analisis sentimen di media sosial terkait Komisi Pemilihan Umum (KPU) sebelum dan sesudah Pilpres 2024 di Indonesia. Populasi sampel yang relevan adalah pengguna aktif platform X yang membicarakan topik terkait KPU dan Pilpres 2024, dengan menggunakan data teks yang mengandung kata kunci KPU sebagai jenis data yang akan dianalisis dalam penelitian ini.

II. METODE PENELITIAN

Metode decision tree cenderung efisien dalam hal waktu komputasi karena strukturnya yang terorganisir secara hirarkis. Ini membuatnya dapat menangani data besar dengan baik, terutama ketika diterapkan dengan implementasi algoritma yang efisien contohnya seperti CART (Classification and Regression Trees) atau C4.5 [9]. Dalam penelitian ini, fokus utama adalah untuk mengevaluasi pendapat masyarakat terkait Komisi Pemilihan Umum (KPU) dengan menggunakan metode *Decision Tree* untuk mengklasifikasikan sentimen menjadi positif, negatif, dan netral. Tujuan dari pendekatan ini adalah untuk meningkatkan keandalan model serta mencegah terjadinya overfitting atau underfitting. Analisis sentimen diimplementasikan melalui pengolahan data dengan memanfaatkan *library textblob*. Data yang digunakan berasal dari Twitter (melalui aplikasi X) dengan menggunakan kata kunci 'KPU' sebelum dan sesudah pemilihan presiden 2024. Penelitian ini melibatkan langkah-langkah praktis, mulai dari pengumpulan data relevan, pembersihan dan tokenisasi teks, hingga validasi silang untuk mengoptimalkan parameter model. Kemudian performa model akan dievaluasi menggunakan metrik seperti akurasi, presisi, recall, dan F1-score untuk memilih kombinasi parameter terbaik. Setelah melatih dan menguji model, hasil analisis memberikan pemahaman yang mendalam tentang pola sentimen publik terhadap KPU sebelum dan sesudah pemilu, serta faktor-faktor yang memengaruhi keputusan model Decision Tree. Selain menggunakan akurasi, metrik evaluasi seperti presisi, recall, F1-score, juga digunakan untuk mengevaluasi kinerja model pada dataset tidak seimbang. Gambar 1 berikut ini memperlihatkan proses penelitian ini.



Gambar 1 Tahapan Penelitian

A. Crawling Data

Untuk mengumpulkan data dari Twitter, dilakukan pencarian dengan kata kunci 'KPU' pada dua rentang waktu berbeda, yaitu dari 30 Januari 2024 hingga 13 Februari 2024, dan dari 15 Februari 2024 hingga 29 Februari 2024. Proses pengambilan data pertama kali memerlukan instalasi Node.js, diikuti dengan inisialisasi variabel untuk menyimpan tweet yang telah diambil serta menetapkan batasan jumlah tweet yang akan diambil. Setelah itu, data diambil menggunakan metode crawling dengan alat tweet-harvest. Hasilnya disimpan dalam format CSV yang kemudian dibaca ke dalam dataframe menggunakan library pandas untuk analisis lebih mendalam.

B. Preprocessing Data

Preprocessing adalah tahapan yang dirancang untuk menghilangkan noise pada data [10]. Proses ini juga bertujuan untuk meningkatkan mutu data serta menghasilkan pembentukan teks yang lebih terorganisir. Berikut beberapa langkah dalam *preprocessing* [11]:

- Cleaning* : menghilangkan bagian-bagian yang tidak relevan dari data.
- Normalization* : menyesuaikan atau mengubah data menjadi format standar atau format yang lebih mudah untuk diproses.
- Stopwords* : menghapus kata-kata yang tidak memiliki makna signifikan, seperti kata penghubung dan kata ganti orang ketiga.
- Tokenization* : pemisahan teks berdasarkan spasi di antara kata-kata dalam kalimat, di mana setiap kata dianggap sebagai token atau entitas terpisah..
- Stemming* : perubahan kata-kata yang mengandung akar kata atau imbuhan menjadi kata dasar.

C. Translating

Dalam konteks analisis sentimen, tahap translating merujuk pada proses mengubah teks yang awalnya ditulis dalam bahasa asing menjadi bahasa Inggris. Hal ini diperlukan sebagai persyaratan penggunaan *TextBlob*, sebuah library dalam pemrosesan bahasa alami yang mendukung analisis sentimen dalam bahasa Inggris.

D. Labeling

Labeling data merupakan langkah di mana setiap data dalam suatu dataset diberi tanda atau label yang mencerminkan informasi mengenai kategori atau kelas data tersebut, seperti positif, negatif, atau netral. Penentuan label dilakukan berdasarkan nilai polarity score yang ditemukan, dengan kriteria sebagai berikut:

- Nilai polaritas > 0 maka label = positif,
- Nilai polaritas $= 0$ maka label = netral, &
- Nilai polaritas < 0 maka label = negatif

E. Classification

Decision Tree adalah metode prediksi yang memanfaatkan representasi pohon keputusan yang digunakan untuk membuat keputusan dengan mempertimbangkan faktor-faktor yang relevan dalam suatu permasalahan [12]. Berikut adalah langkah-langkah dalam membuat pohon keputusan [13] :

- a. Pada *Decision Tree*, pemilihan atribut untuk menjadi akar pohon dilakukan berdasarkan nilai gain tertinggi yang dihasilkan dari atribut yang sudah ada sebelumnya. Rumus persamaan yang digunakan untuk menghitung nilai gain dapat ditemukan di bawah ini.

$$Gain (S,A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy (S_i) \tag{1}$$

Keterangan :

- A : Atribut
- N : Banyaknya pembagian atribut A
- S : Kumpulan kasus
- |S_i| : Jumlah kasus partisi ke-i
- |S| : Jumlah kasus S.

Berikut adalah persamaan yang digunakan untuk menghitung nilai entropy :

$$Entropy (s) = \sum_{i=1}^n - p_i * Log_2 p_i \tag{2}$$

Keterangan:

- S : Kumpulan kasus
- A : Atribut
- N : Total
- p_i : Rasio S_i terhadap S.

- b. Membuat cabang pada tiap-tiap nilai.
- c. Membagi setiap kasus ke cabang.
- d. Membagi kasus pada cabang.
- e. Mengulangi setiap alur agar mendapat kelompok yang serupa dari tiap-tiap divisi maupun kasus.

F. Model Evaluation

Dalam proses evaluasi model, confusion matrix digunakan sebagai instrumen untuk mengukur performa model saat menguji parameter seperti akurasi, presisi, recall, dan skor f1.

III. HASIL DAN PEMBAHASAN

A. Hasil Crawling Data

Data dari platform Twitter atau X dicari menggunakan kata kunci 'KPU' sebelum Pemilihan Presiden 2024 dalam rentang waktu 30/02/2024 hingga 13/02/2024. Dataset setelah pemilihan presiden 2024 diperoleh dari rentang waktu 15/02/2024 hingga 29/022024. Deskripsi hasil pencarian ini kemudian disajikan dalam Tabel 1.

Tabel 1 Hasil Crawling Data twitter

Kata Kunci	Tanggal	Jumlah Tweet
KPU	30/01/2024 - 13/01/2024	2900
KPU	15/02/2024 - 29/02/2024	2894
Total Tweet		5794

Tabel 2 Hasil *Cleaning*

Pra Pilpres	Pasca Pilpres	Cleaning Pra Pilpres	Cleaning Pasca Pilpres
VIVAcoid 🙄🙄🙄 mengingatkan ancaman?? selama ini kemana aja sudah terlalu banyak pelanggaran kalian melempem aja. lembaga negara seperti ini mesti dilebur jd 1 dgn kpu aja, pemborosan apbn gak tau target kerja dan outputnya apaan	TeddGus Bacot aja gede. Hasil resmi kpu aja blom selesai... Trus partai loe lolos gak? 🙄🙄🙄🙄🙄🙄🙄🙄	mengingatkan ancaman selama ini kemana aja sudah terlalu banyak pelanggaran kalian melempem aja lembaga negara seperti ini mesti dilebur jd dgn kpu aja pemborosan apbn gak tau target kerja dan outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai trus partai loe lolos gak
rmol_id Yg bantah mah harusnya MK, KPU, Banwaslu, Kementerian Dalam Negeri, dll. Apa hubungannya sama Alumni UI 🙄🙄🙄	RagilSemar Paslon 01 segera laporkan ribuan temuan kecurangan BAWASLU /KPU sebut blm ada kecurangan buka lebar ² matamu kecurangan tersebar di 38 Provinsi 🙄🙄🙄	yg bantah mah harusnya mk kpu banwaslu kementerian dalam negeri dll apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm ada kecurangan buka lebar ² matamu kecurangan tersebar di provinsi

2. Hasil Normalization

Tahap *normalization* adalah proses mengubah data teks menjadi format standar yang konsisten dan dapat diproses dengan lebih mudah. Ini melibatkan pemrosesan seperti mengubah kata-kata typo atau disingkat menjadi bentuk yang benar dan mengubah kata-kata menjadi bentuk dasarnya. Misalnya, mengubah "yg" menjadi "yang" dan "dr" menjadi "dari". Tujuan normalisasi adalah untuk menyederhanakan teks sehingga lebih mudah dianalisis dan dipahami oleh algoritma. Hasil *normalization* diuraikan dalam Tabel 3 berikut ini.

Tabel 3 Hasil *normalization*

Cleaning Pra Pilpres	Cleaning Pasca Pilpres	Normalization Pra Pilpres	Normalization Pasca Pilpres
mengingatkan ancaman selama ini kemana aja sudah terlalu banyak pelanggaran kalian melempem aja lembaga negara seperti ini mesti dilebur jd dgn kpu aja pemborosan apbn gak tau target kerja dan outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai trus partai loe lolos gak	mengingatkan ancaman selama ini kemana aja sudah terlalu banyak pelanggaran kalian melempem aja lembaga negara seperti ini mesti dilebur jadi dengan kpu aja pemborosan apbn gak tahu target kerja dan outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai terus partai loe lolos gak
yg bantah mah harusnya mk kpu banwaslu kementerian dalam negeri dll apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm ada kecurangan buka lebar ² matamu kecurangan tersebar di provinsi	yang bantah mah harusnya mk kpu banwaslu kementerian dalam negeri dll apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm ada kecurangan buka lebar ² matamu kecurangan tersebar di provinsi

3. Hasil Stopwords

Pada tahap *stopwords*, kata-kata umum yang tidak memiliki kontribusi signifikan terhadap makna konten dieliminasi dari teks dalam proses analisis teks. Ini termasuk kata-kata seperti "yang", "dan", "atau", "di", "ke", "itu" dan "pada" dan lainnya yang dianggap tidak relevan. Tujuan utamanya adalah meningkatkan akurasi analisis dengan memprioritaskan kata-kata kunci yang lebih relevan dalam memahami teks. Hasil *stopwords* diperlihatkan pada Tabel 4 dibawah ini.

Tabel 4 Hasil *stopwords*

Normalization Pra Pilpres	Normalization Pasca Pilpres	Stopwords Pra Pilpres	Stopwords Pasca Pilpres
mengingatkan ancaman selama ini kemana aja sudah terlalu banyak pelanggaran kalian melempem aja lembaga negara seperti ini mesti dilebur jadi dengan kpu aja pemborosan apbn gak tahu target kerja dan outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai terus partai loe lolos gak	mengingatkan ancaman selama kemana aja terlalu banyak pelanggaran kalian melempem aja lembaga negara ini mesti dilebur jadi kpu aja pemborosan apbn gak tahu target kerja outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai terus partai loe lolos gak
yang bantah mah harusnya mk kpu banwaslu kementerian dalam negeri dll apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm ada kecurangan buka lebar ² matamu kecurangan tersebar di provinsi	bantah mah harusnya mk kpu banwaslu kementerian negeri apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm kecurangan buka lebar ² matamu kecurangan tersebar provinsi

4. Hasil Tokenization

Sebelum dianalisis lebih lanjut, teks dioperasikan untuk dibagi menjadi token, yang dapat berupa kalimat, huruf, ataupun kata. Tokenization adalah entitas seperti kata, angka, simbol, tanda baca, dan lainnya yang dapat diidentifikasi secara terpisah [15]. *Tokenization* memecah teks menjadi bagian-bagian kecil yang disebut token, seperti kata-kata atau frasa pendek. Tujuannya adalah untuk memfasilitasi analisis dan pemrosesan teks lebih lanjut dengan mengidentifikasi unit-unit dasar yang dimaksudkan. Hasil *tokenization* dapat diamati pada Tabel 5 dibawah ini.

Tabel 5 Hasil *tokenization*

Stopwords Pra Pilpres	Stopwords Pasca Pilpres	Tokenization Pra Pilpres	Tokenization Pasca Pilpres
mengingatkan ancaman selama kemana aja terlalu banyak pelanggaran kalian melempem aja lembaga negara ini mesti dilebur jadi kpu aja pemborosan apbn gak tahu target kerja outputnya apaan	bacot aja gede hasil resmi kpu aja blom selesai terus partai loe lolos gak	mengingatkan,ancaman, selama,kemana,aja,terlalu, banyak,pelanggaran,kalian, melempem,aja,lembaga, negara,ini,mesti,dilebur, jadi,kpu,aja,pemborosan, apbn,gak,tahu,target,kerja, outputnya,apaan	bacot,aja,gede,hasil,resmi, kpu,aja,blom,selesai,terus, partai,loe,lolos,gak
bantah mah harusnya mk kpu banwaslu kementerian negeri apa hubungannya sama alumni ui	paslon segera laporkan ribuan temuan kecurangan bawaslu kpu sebut blm kecurangan buka lebar ² matamu kecurangan tersebar provinsi	bantah,mah,harusnya,mk, kpu,banwaslu,kementerian, negeri,apa,hubungannya, sama,alumni,ui	paslon,segera,laporkan, ribuan,temuan,kecurangan, bawaslu,kpu,sebut,blm, kecurangan,buka,lebar ² , matamu,kecurangan, tersebar,provinsi

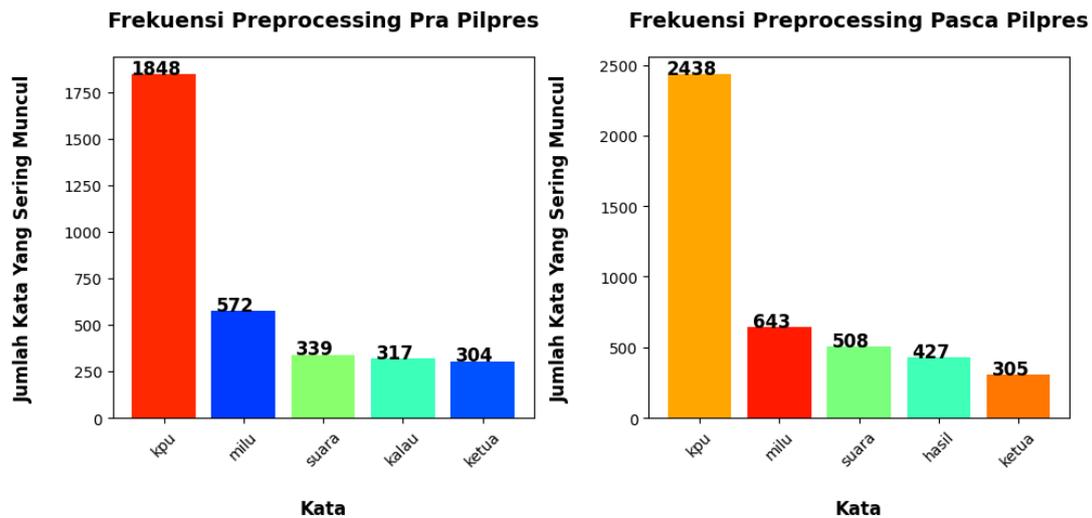
5. Hasil Stemming

Tahap *stemming* melibatkan pengubahan kata-kata berimbuhan menjadi kata dasar menggunakan algoritma khusus [16]. Contohnya, kata "mengingatkan" akan diubah menjadi "ingat", dan "kecurangan" akan menjadi "curang". Tujuannya adalah untuk menyederhanakan kata-kata sehingga variasi morfologis dari kata-kata yang sama dianggap sebagai satu entitas. Hasil *stemming* dapat diperhatikan pada Tabel 6 berikut.

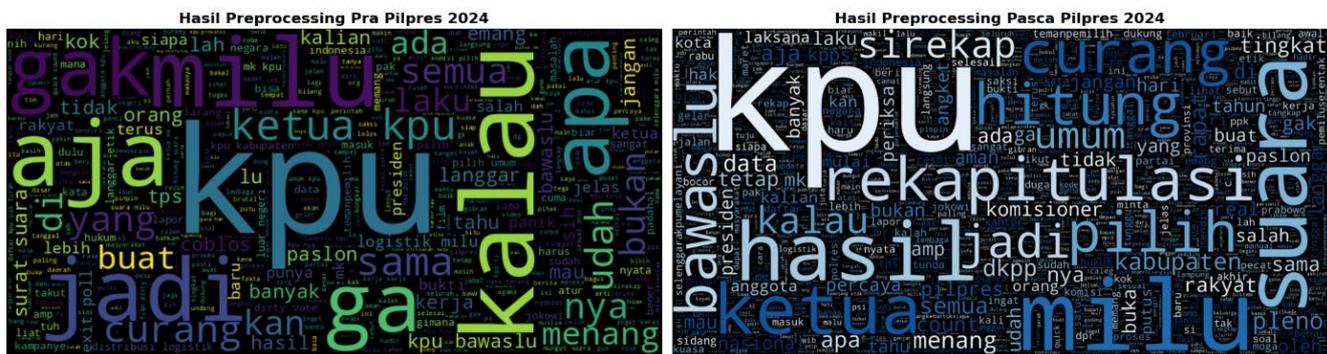
Tabel 6 Hasil *stemming*

Tokenization Pra Pilpres	Tokenization Pasca Pilpres	Stemming Pra Pilpres	Stemming Pasca Pilpres
mengingatkan,ancaman, selama,kemana,aja,terlalu, banyak,pelanggaran,kalian, melempem,aja,lembaga, negara,ini,mesti,dilebur, jadi,kpu,aja,pemborosan, apbn,gak,tahu,target,kerja, outputnya,apaan	bacot,aja,gede,hasil,resmi, kpu,aja,blom,selesai,terus, partai,loe,lolos,gak	ingat ancam lama mana aja terlalu banyak langgar kalian melempem aja lembaga negara ini mesti lebur jadi kpu aja boros apbn gak tahu target kerja outputnya apa	bacot aja gede hasil resmi kpu aja blom selesai terus partai loe lolos gak
bantah,mah,harusnya,mk, kpu,banwaslu,kementerian, negeri,apa,hubungannya, sama,alumni,ui	paslon,segera,laporkan, ribuan,temuan,kecurangan, bawaslu,kpu,sebut,blm, kecurangan,buka,lebar ² , matamu,kecurangan, tersebar,provinsi	bantah mah harus mk kpu banwaslu menteri negeri apa hubung sama alumni ui	paslon segera lapor ribu temu curang bawaslu kpu sebut blm curang buka lebar mata curang sebar provinsi

Setelah tahapan *preprocessing data* selesai, termasuk *cleaning*, *normalization*, *stopwords*, *tokenization* dan *stemming*, *visualisasi* seperti *word cloud* dan *plot* dapat memberikan gambaran tentang frekuensi kata-kata. Dengan visualisasi ini, pola dan tren dalam data dapat dengan mudah terlihat. Deskripsi visualisasi data dapat dicermati dalam Gambar 4 (plot) dan Gambar 5 (wordcloud).



Gambar 4 Hasil Preprocessing (plot)

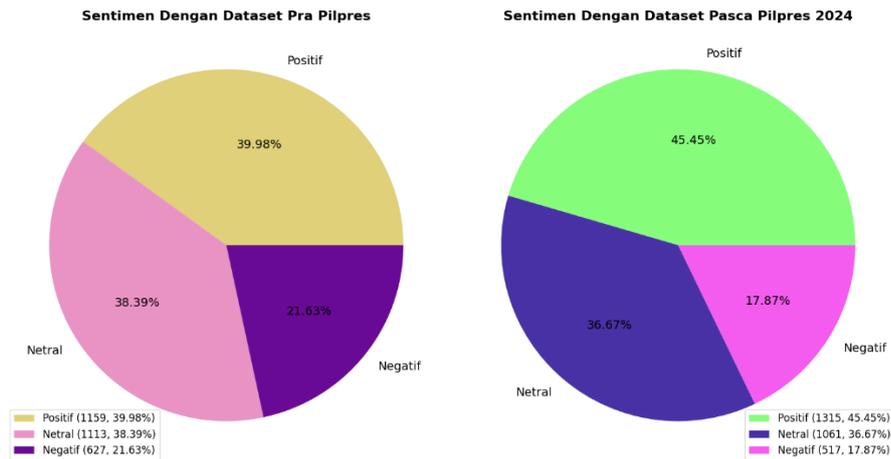


Gambar 5 Hasil Preprocessing (wordcloud)

C. Analisis Sentimen

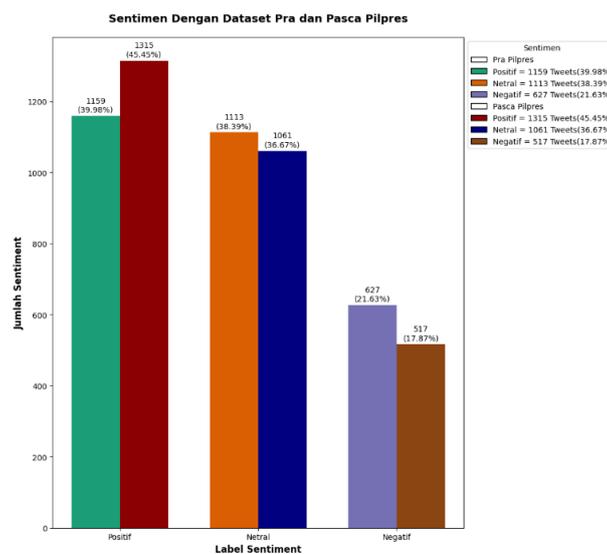
Analisis sentimen melibatkan pengklasifikasian polaritas teks dalam dokumen atau kalimat untuk menentukan apakah pendapat tersebut bersifat positif, netral atau negatif [17]. Analisis sentimen telah menjadi fokus utama penelitian dalam ilmu komputer, terutama seiring dengan meluasnya informasi di internet. Beberapa istilah yang sering digunakan untuk merujuk pada studi ini meliputi analisis sentimen, penambangan opini, ekstraksi opini, dan penambangan sentimen, yang semuanya tergabung dalam ranah analisis sentimen. Konsep analisis sentimen sering dianggap setara dengan opinion mining karena keduanya berfokus pada identifikasi pendapat yang positif atau negatif. [18].

Dalam hasil penelitian ini, ada perubahan sentimen yang signifikan terhadap Komisi Pemilihan Umum (KPU) sebelum dan setelah Pemilihan Presiden 2024 berdasarkan analisis data yang disajikan. Sebelum Pilpres 2024, sentimen positif sebesar 39.98%, netral 38.39%, dan negatif 21.63%. Setelah Pilpres 2024, sentimen positif meningkat menjadi 45.45%, sedangkan netral turun menjadi 36.67%, dan negatif juga turun menjadi 17.87%. Informasi tambahan ada pada gambar 6 dan 7. Uraian lebih lanjut tersedia di bawah ini.



Gambar 6 Hasil analisis sentimen Pra Pilpres dan Pasca Pilpres 2024

Sentimen positif meningkat secara signifikan dari 39.98% menjadi 45.45%, menandakan bahwa mayoritas pengguna platform media sosial X atau Twitter mengurangi pendapat buruk terhadap Komisi Pemilihan Umum (KPU). Sebaliknya, terjadi penurunan yang cukup besar pada sentimen negatif dari 21.63% menjadi 17.87%, mengindikasikan adanya perubahan yang lebih positif dalam pandangan masyarakat terhadap KPU, mungkin dipengaruhi oleh situasi politik saat ini. Selain itu, sentimen netral juga mengalami penurunan dari 38.39% menjadi 36.67%, menunjukkan perubahan dalam penilaian terhadap kinerja KPU setelah pemilihan presiden 2024, yang mungkin disebabkan oleh isu-isu politik atau kebijakan pemerintah.



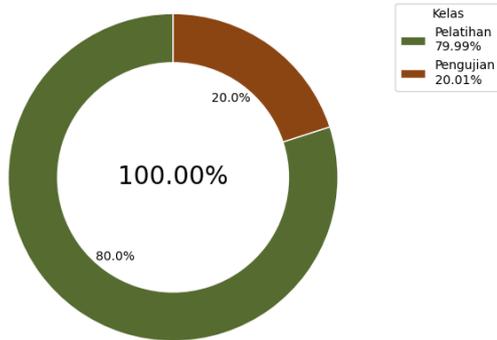
Gambar 7 Persentase hasil analisis sentimen Pra Pilpres dan Pasca Pilpres 2024

Pemerintah harus merespons isu-isu sensitif di media sosial secara bijaksana dan tepat waktu untuk menghindari dampak negatif. Untuk memberikan pemahaman yang lebih baik kepada masyarakat dan mengurangi ketidakpastian, penting untuk meningkatkan tingkat transparansi dalam komunikasi politik. Evaluasi mendalam terhadap kebijakan yang kurang populer juga diperlukan untuk perbaikan. Pemerintah harus terus memperkuat kinerja dan pencapaian yang dapat ditunjukkan secara transparan, termasuk dalam pembangunan dan program-program positif.

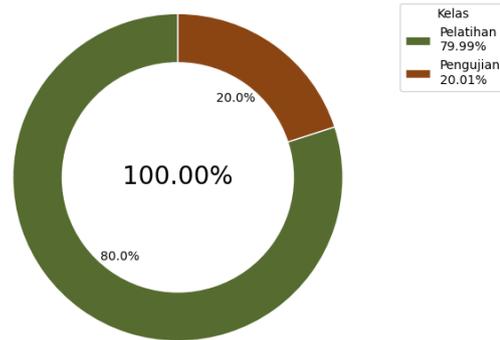
D. Evaluasi Model

Dataset terbagi menjadi 2 bagian sebelum melatih model, 80% untuk data pelatihan dan 20% untuk pengujian. Pembagian data ini kemudian divisualisasikan dalam gambar 8 berikut ini.

Persentase Data Uji dan Training Pra Pilpres 2024

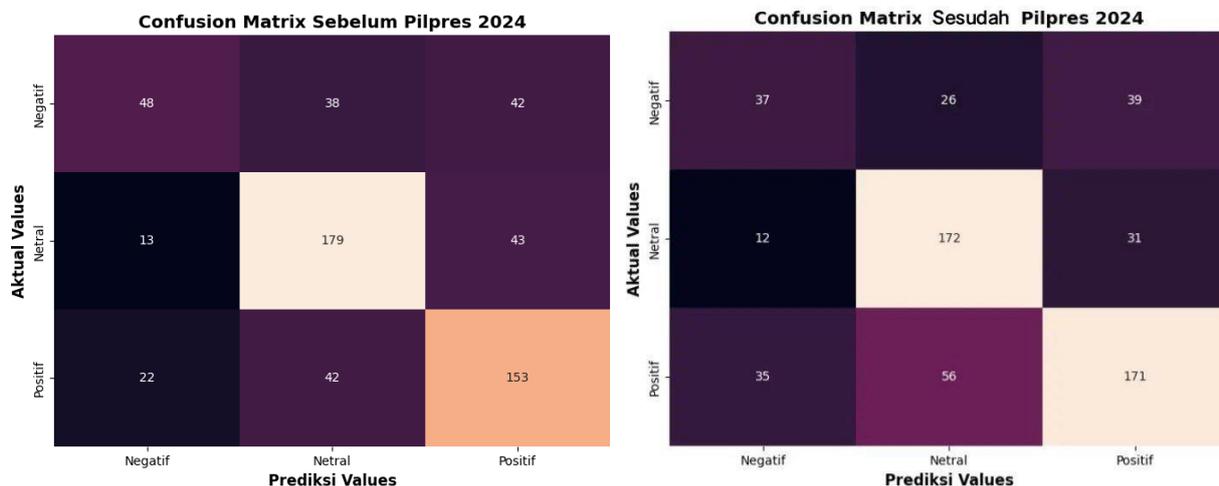


Persentase Data Uji dan Training Pasca Pilpres 2024



Gambar 8 Pembagian dataset pelatihan dan pengujian

Dalam melatih model *Decision Tree*, evaluasi hasil merupakan langkah berikutnya setelah data pelatihan dan pengujian dibagi dengan proporsi yang tepat. Proses evaluasi dilakukan dengan menguji model menggunakan dataset pengujian. Kemudian hasil pengujian digambarkan dalam bentuk *confusion matrix* [19], seperti yang terlihat dalam gambar 9. Dengan pembagian data yang telah dilakukan, langkah evaluasi ini dapat memberikan gambaran akurat tentang kinerja model *Decision Tree* yang telah dilatih.



Gambar 9 Confusion Matrix KPU Pra Pilpres dan Pasca Pilpres 2024

Acuan dalam menentukan nilai parameter seperti precision, recall, dan f1-score adalah *confusion matrix* yang tercantum dalam gambar 9 di atas. Berdasarkan dataset Pra Pilpres, nilai precision, recall dan f1-score tertinggi diperoleh oleh sentimen netral. Akurasi prediksi secara keseluruhan adalah sebesar 63. Sentimen positif memperoleh nilai precision tertinggi sedangkan sentimen netral mencapai nilai tertinggi pada parameter recall dan f1-score, sementara nilai accuracy prediksi sebesar 67 untuk dataset Pasca Pilpres. Perbandingan hasil prediksi model dapat diamati dalam tabel 7 dibawah ini.

Tabel 7 Perbandingan Hasil prediksi model

	Pra			Pasca		
	precision	recall	f1-score	precision	recall	f1-score
negatif	57	37	45	46	38	42
netral	66	73	69	70	80	75
positif	63	69	66	71	68	69

Model dataset KPU setelah Pemilihan Presiden 2024 memperoleh nilai akurasi tertinggi secara keseluruhan. Selain itu, parameter recall dan f1-score juga mencatatkan nilai tertinggi pada dataset yang sama, sementara parameter precision juga memperoleh nilai tertinggi pada dataset tersebut.

Metode Decision Tree memiliki kelebihan dan kelemahan. Kelebihannya adalah kemudahan untuk dipahami dan cocok digunakan dengan data besar, namun kelemahannya termasuk rentan terhadap overfitting dan sulit menangani hubungan antar informasi yang kompleks [20]. Hasil penelitian ini penting karena memberikan

gambaran tentang pandangan masyarakat terhadap KPU. Dengan mengetahui kata-kata atau topik yang memengaruhi sentimen positif atau negatif terhadap KPU, lembaga dapat mengetahui masalah utama yang perlu diatasi untuk meningkatkan dukungan masyarakat. Namun, analisis dari media sosial perlu diperlakukan dengan hati-hati karena bisa dipengaruhi oleh cara pengguna platform tersebut, dan mungkin tidak mewakili pandangan semua orang.

IV. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang sudah dilakukan, analisis data menunjukkan adanya perubahan sentimen yang signifikan terhadap Komisi Pemilihan Umum (KPU) sebelum dan sesudah Pemilihan Presiden 2024. Sebelum pemilihan tersebut, persentase sentimen positif sebesar 39.98%, sedangkan sentimen netral dan negatif masing-masing adalah 38.39% dan 21.63%. Namun, setelah Pemilihan Presiden 2024, terjadi peningkatan signifikan dalam sentimen positif menjadi 45.45%, sementara sentimen netral dan negatif masing-masing mengalami penurunan menjadi 36.67% dan 17.87%. Model ini mencapai akurasi tertinggi sebesar 67% pada dataset setelah Pilpres 2024, menunjukkan performa yang baik dalam memprediksi sentimen terkait KPU. Untuk penelitian selanjutnya mengenai Komisi Pemilihan Umum (KPU), disarankan untuk menggunakan dataset yang jumlah lebih besar, memperluas sumber data selain dari Twitter, seperti platform media sosial lainnya, dan juga mempertimbangkan penggunaan metode klasifikasi lain seperti Naive Bayes.

DAFTAR PUSTAKA

- [1] L. A. Andika, P. A. N. Azizah, and R. Respatiwiulan, "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," *Indones. J. Appl. Stat.*, vol. 2, no. 1, p. 34, 2019, doi: 10.13057/ijas.v2i1.29998.
- [2] E. Febriyani and H. Februariyanti, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Di Twitter," *J. TEKNO KOMPAK*, vol. 17, no. 2, pp. 25–38, 2022, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/2061>
- [3] S. Suryono and E. Taufiq Luthfi, "Analisis sentimen pada Twitter dengan menggunakan metode Naive Bayes Classifier," *Jnanaloka*, pp. 81–86, 2021, doi: 10.36802/jnanaloka.2020.v1-no2-81-86.
- [4] A. Risma Alifvia, U. Saprudin, S. Dharma Wacana Metro, J. Kenanga No, K. Metro Barat, and K. Metro, "Analisis Sentimen Review Data Twitter Komisi Pemilihan Umum (Kpu) Menggunakan Metode Naive Bayes," *J. Inf. dan Komput.*, vol. 11, no. 1, p. 2023, 2023, [Online]. Available: <https://ojs.dccokotabumi.ac.id/index.php/jik/article/view/407>
- [5] L. D. Mahbubah and E. Zuliarso, "Analisa Sentimen Twitter Pada Pilpres 2019 Menggunakan Algoritma Naive Bayes," *Sintak*, pp. 194–195, 2019, [Online]. Available: <https://www.unisbank.ac.id/ojs/index.php/sintak/article/view/7585/2558>
- [6] Febriansyah Putra and H. Patra, "Analisis Hoax pada Pemilu: Tinjauan dari Perspektif Pendidikan Politik," *Naradidik J. Educ. Pedagog.*, vol. 2, no. 1, pp. 95–102, 2023, doi: 10.24036/nara.v2i1.119.
- [7] A. Irma Purnamasari and A. Rinaldi Dikananda, "Klasifikasi Kualitas Berita Pada Majalah Menggunakan Metode Decision Tree," *J. Teknol. Ilmu Komput.*, vol. 1, no. 2, pp. 48–54, 2023, doi: 10.56854/jtik.v1i2.52.
- [8] D. P. Ray, F. N. Hasan, and A. R. Dzirkillah, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Terhadap KPU 2024 Berdasarkan Tweet Media Sosial Twitter Menggunakan Algoritma Naive Bayes," *Media Online*, vol. 4, no. 4, pp. 2235–2243, 2024, doi: 10.30865/klik.v4i4.1587.
- [9] D. P. Sukma, S. Defit, and G. W. Nurcahyo, "Identifikasi Tingkat Kerusakan Peralatan Labor Teknik Komputer Jaringan Menggunakan Metode Decision Tree," *J. Sistim Inf. dan Teknol.*, vol. 3, pp. 275–280, 2021, doi: 10.37034/jsisfotek.v3i4.78.
- [10] J. A. Septian, T. M. Fachrudin, and A. Nugroho, "Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor," *J. Intell. Syst. Comput.*, vol. 1, no. 1, pp. 43–49, 2019, doi: 10.52985/insyst.v1i1.36.
- [11] L. A. Pramesti and N. Pratiwi, "Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1145–1154, 2023, doi: 10.47065/josh.v4i4.3807.
- [12] Y. A. Singgalen, "Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine," *Media Online*, vol. 4, no. 1, pp. 323–332, 2023, doi: 10.30865/klik.v4i1.1153.
- [13] T. A. Q. Putri, A. Triayudi, and R. T. Aldisa, "Implementasi Algoritma Decision Tree dan Naive Bayes Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks," *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 641–649, 2023, doi: 10.47065/josh.v4i2.2949.
- [14] Rina Noviana and Isram Rasal, "Penerapan Algoritma Naive Bayes Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter," *J. Tek. dan Sci.*, vol. 2, no. 2, pp. 51–60, 2023, doi: 10.56127/jts.v2i2.791.
- [15] J. Florensus Sianipar, Y. R. Ramadhan, and I. Jaelani, "Kajian Ilmiah Informatika dan Komputer," *Media Online*, vol. 4, no. 1, pp. 360–367, 2023, doi: 10.30865/klik.v4i1.1033.
- [16] M. Z. Yumarlin, "Penerapan Metode K-Nearest Neighbor Dan Decision Tree Untuk Analisis Sentimen (Studi Kasus Mario Dandi)," *IJITECH Indones. J. Inf.*, pp. 1–6, 2023, [Online]. Available: <https://ojsnu.nuonorogo.or.id/index.php/ijitech/article/view/54%0Ahttps://ojsnu.nuonorogo.or.id/index.php/ijitech/article/download/54/30>
- [17] D. Ramadhan and E. B. Setiawan, "Analisis Sentimen Program Acara di SCTV pada Twitter Menggunakan Metode Naive Bayes dan Support Vector Machine," *... .Telkomuniversity.Ac.Id*, vol. 6, no. 2, pp. 9736–9743, 2019, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/10708>
- [18] M. Arvyantomo and N. Ratama, "Analisis Sentimen Masyarakat Indonesia terhadap Invasi Russia di Ukraina menggunakan Metode Naive Bayes pada Media Sosial Facebook," *Log. J. Ilmu Komput.*, vol. 1, no. 4, pp. 705–717, 2023, [Online]. Available: <https://www.journal.mediapublikasi.id/index.php/logic/article/view/2142>
- [19] S. S. Salim and J. Mayary, "Analisis Sentimen Pengguna Twitter Terhadap Dompot Elektronik Dengan Metode Lexicon Based Dan K – Nearest Neighbor," *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.
- [20] C. A. Sari, A. Sukmawati, R. P. Aprilli, P. S. Kayaningtias, and N. Yudistira, "Perbandingan Metode Naive Bayes, Support Vector Machine Dan Decision Tree Dalam Klasifikasi Konsumsi Obat," *J. Litbang Edusaintech*, vol. 3, no. 1, pp. 33–41, 2022, [Online]. Available: <http://dx.doi.org/10.51402/jle.v3i1.47>