

## ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP KEBIJAKAN DEDI MULYADI DENGAN METODE MACHINE LEARNING (STUDI KASUS : KENAKALAN REMAJA)

Rizky Budi Aprianto, Ade Eviyanti, Hamzah Setiawan, Ika Ratna Indra Astutik

Informatika, Universitas Muhammadiyah Sidoarjo

Jl. Raya Gelam No.250, Sidoarjo, Indonesia

adeeviyanti@umsida.ac.id

### ABSTRAK

Kenakalan remaja merupakan isu sosial yang memunculkan kebijakan kontroversial berupa pengiriman anak bermasalah ke barak militer oleh Dedi Mulyadi. Kebijakan tersebut menimbulkan beragam tanggapan masyarakat di platform YouTube yang mencerminkan dinamika opini publik. Permasalahan dalam penelitian ini adalah bagaimana mengklasifikasikan sentimen komentar masyarakat secara akurat serta mengevaluasi performa metode ensemble machine learning dibandingkan algoritma tunggal. Penelitian ini bertujuan menganalisis kecenderungan opini publik dan mengukur kinerja algoritma klasifikasi dalam memetakan sentimen komentar YouTube. Metode yang digunakan meliputi pengumpulan 7.875 komentar melalui YouTube Data API, pelabelan otomatis menggunakan InSet Lexicon, preprocessing teks, serta pembobotan fitur dengan TF-IDF. Model klasifikasi yang diterapkan adalah Naïve Bayes, K-Nearest Neighbor (KNN), dan Logistic Regression, yang kemudian dikombinasikan menggunakan teknik Ensemble Majority Voting. Hasil penelitian menunjukkan bahwa sentimen publik didominasi oleh kelas positif (52,4%) dibandingkan negatif (47,6%). Logistic Regression memperoleh akurasi tertinggi sebesar 0,84, sedangkan model ensemble menunjukkan performa stabil dengan akurasi sekitar 0,82. Temuan ini menunjukkan bahwa pendekatan ensemble mampu meningkatkan stabilitas klasifikasi sentimen pada isu kebijakan publik di media sosial.

**Kata kunci :** Analisis Sentimen, YouTube, Dedi Mulyadi, Ensemble Machine Learning

### 1. PENDAHULUAN

Permasalahan kenakalan remaja saat ini menjadi topik yang diperbincangkan secara serius di Indonesia karena kompleksitas persoalannya yang mencakup aspek sosial dan psikologis. Kenakalan remaja, atau secara akademis dikenal sebagai juvenile delinquency, merupakan gejala patologis pada anak-anak maupun remaja yang muncul akibat perilaku yang tidak sesuai dengan norma sosial sehingga mengarah pada tindakan menyimpang [1]. Fenomena ini mencakup berbagai perilaku seperti merokok, tawuran, membolos sekolah, hingga keterlibatan dalam pergaulan bebas yang berpotensi memberikan dampak jangka panjang bagi masyarakat [2]. Munculnya kenakalan remaja dipengaruhi oleh interaksi faktor internal dan eksternal, di antaranya krisis identitas dan lemahnya kontrol diri, serta pengaruh teman sebaya, media sosial, dan kondisi keluarga yang kurang harmonis [3][4]. Berbagai penelitian menunjukkan bahwa lingkungan keluarga, pergaulan, dan kondisi sosial memiliki peran signifikan dalam membentuk perilaku remaja [1][2][3][4].

Upaya penanganan melalui pendekatan pendidikan dan pembinaan terus dilakukan guna mengubah perilaku remaja ke arah yang lebih positif [5]. Namun, kegagalan sistem dalam mendidik generasi muda mendorong mencari pendekatan alternatif yang dinilai lebih tegas. Salah satu kebijakan yang menarik perhatian luas berasal dari perintah Gubernur Jawa Barat, Dedi Mulyadi, untuk mengirimkan anak-anak yang dianggap bermasalah ke barak militer guna mengikuti pembentukan karakter

dengan disiplin ketat [6]. Kebijakan ini memunculkan beragam tanggapan dari masyarakat, baik yang mendukung maupun yang menolak, khususnya pada platform digital seperti YouTube. Ribuan komentar yang muncul pada video terkait kebijakan tersebut menunjukkan adanya dinamika opini publik yang kompleks dan beragam. Banyaknya data komentar serta variasi bahasa yang digunakan menjadikan analisis secara manual kurang efektif, sehingga diperlukan pendekatan analisis sentimen untuk mengolah opini publik secara otomatis dan terstruktur [7].

Berdasarkan kondisi tersebut, permasalahan yang dikaji dalam penelitian ini adalah bagaimana mengidentifikasi dan mengklasifikasikan sentimen komentar masyarakat terhadap kebijakan Dedi Mulyadi secara akurat, serta bagaimana mengetahui tingkat akurasi model dalam penggunaan metode ensemble machine learning. Analisis sentimen diperlukan untuk menentukan apakah tanggapan penonton bersifat positif atau negatif terhadap kebijakan tersebut [8]. Selain itu, diperlukan evaluasi terhadap performa beberapa algoritma machine learning guna mengetahui model yang paling optimal dalam mengklasifikasikan sentimen pada data komentar YouTube.

Penelitian ini bertujuan untuk menganalisis kecenderungan opini publik terhadap kebijakan pembinaan remaja oleh Dedi Mulyadi berdasarkan ulasan masyarakat di YouTube, serta menilai tingkat akurasi algoritma machine learning dalam melakukan klasifikasi sentimen. Untuk mencapai tujuan tersebut,

penelitian ini menerapkan teknik pengumpulan data melalui scraping menggunakan YouTube Data API, kemudian dilakukan proses *text preprocessing* yang meliputi *cleansing*, *case folding*, *tokenization*, dan *stopword removal* guna membersihkan serta menyeragamkan data teks [9]. Tahap pembobotan kata menggunakan TF-IDF dilakukan untuk merepresentasikan kepentingan fitur dalam data teks [10]. Data kemudian dilabeli menggunakan pendekatan lexicon sebelum dilakukan pemisahan dataset ke dalam data pelatihan dan pengujian, serta pengukuran kinerja model menggunakan metode ensemble yang menggabungkan beberapa algoritma klasifikasi.

Fokus utama penelitian ini adalah mengevaluasi kemampuan algoritma Naïve Bayes, K-Nearest Neighbor, dan Logistic Regression serta metode Ensemble Majority Voting dalam mengklasifikasikan sentimen komentar masyarakat terhadap kebijakan Dedi Mulyadi pada *platform* YouTube. Pengujian dilakukan menggunakan tiga variasi proporsi data pelatihan dan pengujian, yaitu 70:30, 65:35, dan 60:40, guna menilai stabilitas serta konsistensi performa model pada berbagai skenario pembagian data. Kinerja masing-masing algoritma dianalisis berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui tingkat ketepatan serta keseimbangan hasil klasifikasi. Hasil penelitian ini diharapkan dapat memperkaya kajian analisis sentimen pada isu kebijakan publik di media sosial serta menjadi referensi dalam memahami dinamika opini masyarakat secara lebih objektif dan terukur.

## 2. TINJAUAN PUSTAKA

### 2.1. Analisis Sentimen

Analisis sentimen merupakan teknik yang dimanfaatkan dalam mengetahui kecenderungan persepsi pengguna dengan menganalisis teks, sehingga dapat ditentukan polaritas sentimen yang terkandung, baik positif maupun negatif, terhadap suatu objek tertentu. Dalam konteks komentar YouTube, analisis sentimen digunakan untuk mengevaluasi respons emosional serta opini masyarakat terhadap suatu topik berdasarkan isi komentar yang disampaikan oleh pengguna [11].

### 2.2. Penelitian Terdahulu

Beberapa penelitian terdahulu menerapkan analisis sentimen pada komentar youtube menggunakan algoritma klasifikasi berbasis *machine learning*. Herdiansyah dan Yuliana melakukan analisis sentimen publik terhadap kebijakan Kampus Merdeka berdasarkan komentar pada video YouTube resmi Kemendikbud. Data dikumpulkan melalui proses *crawling* menggunakan google colab dan diklasifikasikan dengan algoritma Naïve Bayes pada sentimen negatif, netral, serta positif. Hasil penelitian menunjukkan dominasi sentimen positif, dengan performa model yang tergolong cukup baik

berdasarkan nilai akurasi, presisi, *recall*, dan F1-score [12].

Selanjutnya, Larasakti dkk. melaksanakan analisis sentiment komentar YouTube pada video podcast Deddy Corbuzier menerapkan metode K-Nearest Neighbor (KNN). Data komentar diproses melalui tahapan preprocessing teks, pembobotan TF-IDF, dan pengukuran kemiripan menggunakan cosine similarity. Hasil pengujian mengindikasikan mengenai model KNN dapat menghasilkan tingkat akurasi secara sangat besar [13].

Penelitian lain oleh Mulyono dan Saprudin mengkaji sentimen komentar YouTube berbahasa Indonesia terkait isu ketenagakerjaan menggunakan algoritma Logistic Regression. Data diperoleh melalui YouTube Data API dan dikelompokkan pada sentimen positif, netral, dan negatif. Tahapan penelitian meliputi *preprocessing* teks, ekstraksi fitur TF-IDF, serta evaluasi model klasifikasi, dengan hasil mengindikasikan mengenai model terdapat kinerja terbaik pada kelas sentimen negatif [14].

Penelitian selanjutnya oleh Fachrudin dkk. menganalisis sentimen pengguna Twitter terhadap kualitas layanan jaringan internet Telkomsel dengan membandingkan algoritma K-Nearest Neighbor (KNN), Support Vector Machine (SVM), serta metode Ensemble KNN-SVM. Data Twitter diklasifikasikan ke dalam sentimen positif, netral, dan negatif melalui tahapan *preprocessing* teks dan pembobotan TF-IDF. Hasil penelitian mengindikasikan bahwa metode *ensemble* KNN-SVM berbasis *average* memberikan performa terbaik dibandingkan model *majority*, serta mengungkap dominasi sentimen negatif terhadap layanan jaringan Telkomsel [15].

Berdasarkan penelitian sebelumnya, berbagai algoritma *machine learning* telah digunakan dalam analisis sentimen media sosial dengan kinerja yang bergantung pada karakteristik data. Namun, penerapan metode *ensemble* pada komentar YouTube terkait kebijakan publik masih terbatas. Dengan demikian, penelitian ini berfokus terhadap penggunaan *ensemble machine learning* untuk menganalisis sentimen komentar YouTube terhadap kebijakan publik.

### 2.3. Algoritma Naïve Bayes

Algoritma Naive Bayes merupakan algoritma pengelompokan berbasis probabilitas dengan bekerja menerapkan Teorema Bayes dengan asumsi independensi antar fitur, sehingga dapat dimanfaatkan dalam memprediksi kelas sebuah data menurut probabilitas kemunculan fitur-fitur [16].

Beberapa penelitian sebelumnya mengindikasikan bahwa algoritma Naïve Bayes mampu memberikan kinerja yang memadai dalam klasifikasi sentimen pada data komentar pengguna di *platform* digital, seperti YouTube, dengan hasil evaluasi yang menunjukkan performa yang cukup baik berdasarkan metrik akurasi, presisi, *recall*, dan F1-score [8][12].

## 2.4. Algoritma K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbor (KNN) merupakan algoritma klasifikasi berbasis *supervised learning* yang menentukan kelas sebuah data menurut mayoritas kelas dari k tetangga terdekatnya, dengan mengukur kedekatan data menggunakan jarak tertentu tanpa asumsi distribusi data [17]. Algoritma ini tidak melakukan asumsi terhadap distribusi data dan menentukan kelas berdasarkan mayoritas kelas dari data tetangga terdekatnya.

Penelitian terkait yang telah dilakukan sebelumnya menunjukkan mengenai algoritma K-Nearest Neighbor (KNN) mampu mencapai kinerja klasifikasi sentimen yang sangat baik pada data komentar YouTube, khususnya ketika diterapkan bersama proses *preprocessing* teks, pembobotan fitur menggunakan TF-IDF, serta pengukuran tingkat kemiripan berbasis *cosine similarity* Larasakti dkk. [11].

## 2.5. Algoritma Logistic Regression

Algoritma Logistic Regression merupakan algoritma klasifikasi berbasis probabilitas dengan menerapkan fungsi logistik dalam memprediksi peluang sebuah data masuk pada kelas tertentu menurut fitur-fitur yang terdapat [18].

Penelitian sebelumnya membuktikan mengenai algoritma Logistic Regression terdapat performa yang cukup baik dalam melakukan klasifikasi sentimen pada komentar YouTube, khususnya ketika didukung oleh tahapan *preprocessing* teks dan pembobotan fitur menggunakan metode TF-IDF, sehingga mampu mengidentifikasi kategori sentimen positif, netral, dan negatif secara efektif [14].

## 2.6. Ensemble Machine Learning

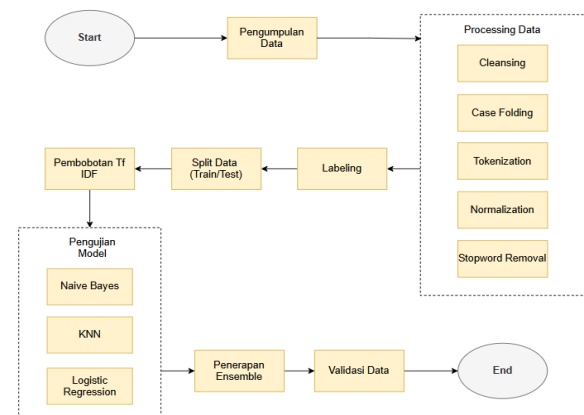
Ensemble Machine Learning merupakan metode dengan mengkombinasikan sejumlah model klasifikasi dalam menghasilkan prediksi secara lebih stabil juga akurat dibandingkan satu model tunggal, dengan memanfaatkan kelebihan tiap algoritma guna meningkatkan kinerja dan mengurangi kesalahan prediksi [19].

Dalam *ensemble machine learning*, salah satu pendekatan yang digunakan adalah metode voting, yaitu menggabungkan prediksi setiap model untuk menentukan kelas paling representatif. Teknik yang diterapkan adalah Hard Voting, yang menetapkan keputusan klasifikasi final melalui penggabungan prediksi setiap algoritma berdasarkan perolehan suara terbanyak [20].

Berdasarkan penelitian terdahulu, dapat diketahui bahwa analisis sentimen pada media sosial telah banyak dilakukan menggunakan beragam algoritma machine learning dengan tingkat kinerja yang bervariasi sesuai karakteristik data. Meskipun demikian, penelitian yang secara khusus mengkaji dan membandingkan penerapan metode *ensemble* pada komentar YouTube terkait isu kebijakan publik masih terbatas. Oleh sebab itu, penelitian ini diarahkan pada

penerapan metode *ensemble machine learning* dalam menganalisis sentimen komentar YouTube terhadap kebijakan publik [15].

## 3. METODE PENELITIAN



Gambar 1. Diagram Tahapan Penelitian

Ada delapan langkah yang dilakukan dalam penelitian ini untuk mendapatkan hasil yang diinginkan, seperti yang terlihat pada Gambar 1.

### 3.1. Sumber dan Pengumpulan Data

Data yang dimanfaatkan untuk penelitian ini adalah data sekunder berupa komentar pengguna pada platform YouTube terkait kebijakan Kang Dedi Mulyadi tentang pengiriman anak bermasalah ke barak militer. Data dikumpulkan dengan memanfaatkan YouTube Data API untuk mengambil komentar secara otomatis dari video yang telah ditentukan, dengan contoh data mentah yang disajikan pada Tabel 1. Komentar berbahasa Indonesia digunakan sebagai data penelitian yang mencerminkan opini masyarakat terhadap kebijakan tersebut. Pengambilan data dilakukan pada periode Juni hingga Desember 2025 dengan total 7.875 komentar, yang selanjutnya disimpan dalam format XLXS untuk dilanjutkan dalam tahapan pengolahan serta analisis data.

Tabel 1. Contoh Data Komentar Youtube

| Video_id    | Author          | Published_at         | Text                       | Likecount |
|-------------|-----------------|----------------------|----------------------------|-----------|
| XxrD9aMNR6E | @uusrusmana3032 | 2025-06-10T16:49:20Z | Manta ppppl KDM            | 0         |
| XxrD9aMNR6E | @uusrusmana3032 | 2025-06-10T16:47:18Z | KDM semoga panjang umurnya | 0         |
| XxrD9aMNR6E | @uusrusmana3032 | 2025-06-10T16:46:30Z | Manta pppp                 | 0         |

### 3.2. Preprocessing Data

Tahap preprocessing dilaksanakan dalam menyiapkan data komentar sebelum permodelan dengan tujuan membersihkan dan menyeragamkan teks agar siap diproses oleh algoritma klasifikasi.

Tahapan yang diterapkan meliputi *case folding* untuk mengubah teks menjadi huruf kecil, *cleaning* untuk menghapus karakter yang tidak relevan, *tokenization* untuk memecah teks menjadi kata, *normalization* untuk menyesuaikan kata tidak baku, serta *stopword removal* untuk menghilangkan kata yang tidak memiliki makna signifikan. Ilustrasi perubahan data teks pada setiap tahapan tersebut disajikan secara detail pada Tabel 2. Secara keseluruhan, rangkaian proses ini bertujuan untuk mengurangi noise dan meningkatkan kualitas fitur teks sehingga dapat diolah secara optimal pada tahap klasifikasi

Tabel 2. Data Hasil Preprocessing

| Tahap            | Komentar                                |
|------------------|-----------------------------------------|
| Data Awal        | KDM semoga panjang umurnya              |
| Case Folding     | kdm semoga panjang umurnya              |
| Cleaning         | kdm semoga panjang umurnya              |
| Tokenization     | ['kdm', 'semoga', 'panjang', 'umurnya'] |
| Normalization    | ['kdm', 'semoga', 'panjang', 'umur']    |
| Stopword Removal | ['kdm', 'semoga', 'umur']               |

### 3.3. Labeling

Setelah tahap preprocessing selesai, data komentar kebijakan Dedi Mulyadi melalui proses pelabelan sentimen otomatis ke dalam kategori positif (label 1) dan negatif (label 0) menggunakan InSet Lexicon Bahasa Indonesia. Mekanisme pelabelan dilakukan dengan mengakumulasi bobot kata berdasarkan kamus sentimen, di mana komentar dengan skor sentimen  $\geq 0$  dikategorikan sebagai positif, sedangkan skor  $< 0$  ditetapkan sebagai negatif. Implementasi teknis dan contoh hasil pelabelan yang menunjukkan keterhubungan antara teks, stopwords removal, skor, hingga label akhir tersebut disajikan secara detail pada Tabel 3.

Tabel 3. Data Hasil Labeling

| Komentar                                         | Stopword removal            | Sentiment Score | Label |
|--------------------------------------------------|-----------------------------|-----------------|-------|
| Isinya daging smua                               | ['isinya', 'daging']        | 0               | 1     |
| Luar biasa pidato yg sangat menginspirasi<br>♥♥♥ | ['pidato', 'menginspirasi'] | 0               | 1     |
| Mantappppl KDM                                   | ['mantap', 'kdm']           | 5               | 1     |

### 3.4. Split Data

Tahap selanjutnya melibatkan pembagian dataset ke dalam dua komponen, diantaranya data latih serta data uji, dalam tiga scenario rasio 70%-30%, 65%-35%, 60%-40%. Data latih dimanfaatkan dalam membangun serta melatih model untuk mengenali pola sentiment, akan tetapi data uji digunakan dalam mengevaluasi kinerja model terhadap data yang belum pernah diproses sebelumnya. Penggunaan beberapa

variasi rasio ini bertujuan untuk menilai kestabilan dan keandalan model dalam menganalisis teks.

### 3.5. Pembobotan TF-IDF

Secara matematis, pembobotan TF-IDF dirumuskan sebagai berikut:

$$TF-IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

dengan:

- $TF(t, d)$  adalah frekuensi kemunculan term  $t$  dalam dokumen  $d$
- $IDF(t)$  mempresentasikan *inverse document frequency* dari term  $t$

Nilai IDF dihitung menggunakan persamaan:

$$IDF(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

Keterangan :

- $N$  adalah total jumlah dokumen
- $df(t)$  adalah jumlah dokumen yang memuat term  $t$

Pada tahap pembobotan fitur, teks pada data latih dan data uji diubah menjadi representasi vektor numerik menggunakan metode TF-IDF. Metode ini memberikan nilai bobot pada setiap kata sesuai tingkat kepentingannya dalam dokumen, sehingga dapat digunakan sebagai fitur input pada algoritma klasifikasi.

### 3.6. Pengujian Model

#### a. K-Nearest Neighbor (KNN)

Algoritma KNN melibatkan pendekatan berbasis kedekatan jarak antar data dalam proses klasifikasi. Penentuan kelas dilakukan dengan menghitung jarak antara data uji dan seluruh data latih, kemudian memilih sejumlah  $k$  data dengan jarak terdekat. Secara matematis, perhitungan jarak dapat dinyatakan dengan rumus Euclidean Distance sebagai berikut:

$$d(x_i, x_j) = \sqrt{\sum (x_{ik} - x_{jk})^2} \quad (3)$$

Keterangan:

- $d(x_i, x_j)$  merupakan jarak antara data ke- $i$  dan data ke- $j$ .
- $x_{ik}$  dan  $x_{jk}$  merupakan nilai fitur ke- $k$  dari masing-masing data.
- $n$  merupakan jumlah fitur.

Setelah diperoleh  $k$  tetangga terdekat, kelas sentimen ditentukan berdasarkan mayoritas kelas dari  $k$  data tersebut.

#### b. Naïve Bayes

Algoritma Naïve Bayes menggunakan pendekatan probabilitas dalam proses klasifikasi dengan mengacu pada Teorema Bayes serta mengasumsikan bahwa setiap atribut bersifat independen terhadap atribut lainnya. Secara matematis, Teorema Bayes dinyatakan sebagai berikut:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (4)$$

Keterangan:

- $P(C|X)$  merupakan probabilitas kelas C dengan kondisi fitur X.
- $P(X|C)$  merupakan probabilitas fitur X pada kondisi kelas C.
- $P(C)$  merupakan probabilitas prior dari kelas C.
- $P(X)$  merupakan probabilitas dari fitur X.

Dalam proses klasifikasi, nilai  $P(X)$  bersifat konstan sehingga tidak memengaruhi perbandingan antar kelas dan dapat diabaikan. Penentuan sentimen dilakukan dengan memilih kelas yang memiliki nilai probabilitas posterior tertinggi.

### c. Logistic Regression

Logistic Regression menggunakan pendekatan fungsi logistik untuk memodelkan probabilitas suatu data termasuk ke dalam kelas tertentu. Proses klasifikasi dilakukan dengan menghitung nilai probabilitas menggunakan fungsi sigmoid. Secara matematis, fungsi tersebut dinyatakan sebagai berikut:

$$P(Y = 1|X) = \frac{1}{(1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)})} \quad (5)$$

Keterangan:

- $P(Y=1|X)$  merupakan probabilitas data termasuk ke dalam kelas tertentu.
- $\beta_0$  merupakan konstanta (intercept).
- $\beta_1, \dots, \beta_n$  merupakan koefisien masing-masing fitur.
- $X_1, \dots, X_n$  merupakan nilai fitur.
- $e$  merupakan bilangan eksponensial.

Nilai probabilitas kemudian dibandingkan dengan ambang batas (*threshold*). Jika melebihi ambang batas, data diklasifikasikan sebagai kelas positif, sedangkan jika di bawahnya diklasifikasikan sebagai kelas negatif.

### 3.7. Penerapan Ensemble

Dalam tahap ini digunakan metode Ensemble Machine Learning dengan menggabungkan model Naïve Bayes, K-Nearest Neighbor (KNN), dan Logistic Regression menggunakan teknik Hard Voting, yaitu penentuan kelas akhir berdasarkan mayoritas hasil prediksi dari masing-masing model. Setiap algoritma menghasilkan prediksi secara independen, kemudian kelas yang paling banyak dipilih ditetapkan sebagai hasil klasifikasi akhir, sehingga memungkinkan peningkatan kinerja model serta perbandingan performa antara metode *ensemble* dan model tunggal.

### 3.8. Validasi Data

Evaluasi kinerja model dilakukan untuk menilai kemampuan dua algoritma dalam mengklasifikasikan sentimen ulasan aplikasi. Proses evaluasi ini

menggunakan *confusion matrix* dan beberapa metrik, yaitu *accuracy*, *precision*, *recall*, serta *F1-score*.

*Accuracy* digunakan untuk mengetahui tingkat ketepatan model dalam melakukan klasifikasi, dengan rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

*Precision* menunjukkan ketepatan model saat memprediksi kelas positif, yang dihitung dengan rumus:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

*Recall* menggambarkan kemampuan model dalam mengidentifikasi data yang benar-benar termasuk kelas positif, dengan rumus:

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

*F1-score* merupakan gabungan dari *precision* dan *recall* dalam bentuk rata-rata harmonis, yang dirumuskan sebagai:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

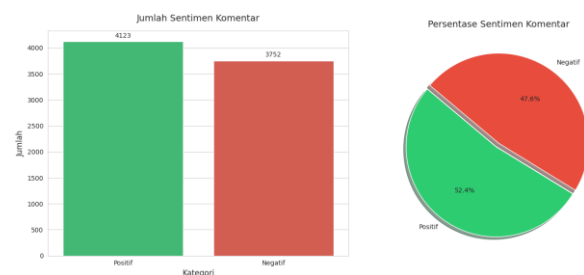
Keterangan :

- TP = prediksi yang benar-benar sesuai dengan kelas positif.
- TN = prediksi yang benar-benar sesuai dengan kelas negatif.
- FP = kesalahan memprediksi negatif sebagai positif.
- FN = kesalahan memprediksi positif sebagai negatif

## 4. HASIL DAN PEMBAHASAN

### 4.1. Deskripsi Data dan Hasil Pelabelan

Dataset sebanyak 7.875 ulasan periode Juni-Desember 2025 dikumpulkan melalui YouTube Data API dan dilabeli secara otomatis menggunakan InSet Lexicon. Hasil pelabelan menunjukkan dominasi sentimen positif sebesar 52,4% (4.123 ulasan) dibandingkan sentimen negatif sebesar 47,6% (3.752 ulasan) sebagaimana ditampilkan pada Gambar 2. Meskipun opini publik kini cenderung lebih mendukung kebijakan Dedi Mulyadi, distribusi ini tetap memiliki keseimbangan kelas yang baik untuk menunjang akurasi pengujian model ensemble machine learning.

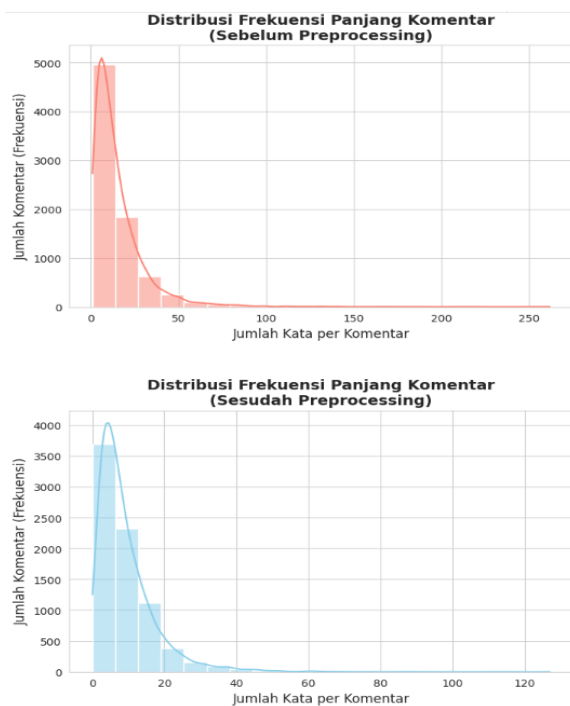


Gambar 2. Distribusi Sentimen

#### 4.2. Implementasi Preprocessing dan Ekstraksi Fitur

Tahap ini mentransformasi teks mentah melalui rangkaian preprocessing—mulai dari cleansing hingga stopword removal—untuk mengeliminasi noise. Efektivitas pembersihan ini dibuktikan pada Gambar 3, di mana terjadi penyusutan jumlah kata yang signifikan setelah elemen non-informatif dihapus.

Setelah data dinyatakan bersih, fitur teks dikonversi menjadi representasi numerik menggunakan metode TF-IDF. Pemberian bobot pada setiap kata berdasarkan tingkat kepentingannya ini bertujuan untuk menghasilkan input yang akurat bagi model ensemble machine learning dalam mengenali pola sentimen terkait kebijakan Dedi Mulyadi secara optimal



Gambar 3. Distribusi Panjang Komentar Sebelum dan Sesudah Preprocessing

### 4.3. Wordcloud

Berdasarkan Gambar 4, terlihat transformasi data yang signifikan di mana bagian atas (sebelum *preprocessing*) didominasi oleh *noise* seperti "yg" dan "dan", sedangkan bagian bawah (sesudah *preprocessing*) berhasil menonjolkan kata kunci substansial seperti "kdm", "bagus", dan "kpai". Penghilangan kata-kata non-informatif ini membuktikan keberhasilan tahap *preprocessing* dalam mengekstraksi fitur teks yang relevan dan bermakna, sehingga siap digunakan untuk proses klasifikasi sentimen yang lebih akurat.



Gambar 4. Wordcloud Sebelum dan Sesudah Preprocessing

#### 4.4. Hasil Pengujian Model

Pemodelan klasifikasi ini memanfaatkan 3 model algoritma Naïve Bayes, K-Nearest Neighbor (KNN), dan Logistic Regression melalui tiga skenario rasio pembagian data, yakni 70:30, 65:35, dan 60:40 sebagaimana disajikan pada Tabel 4. Berdasarkan hasil eksperimen tersebut, Logistic Regression secara konsisten menghasilkan akurasi tertinggi sebesar 0,84. di seluruh skenario pengujian. Di sisi lain, Naïve Bayes menunjukkan stabilitas performa dengan nilai akurasi tetap di angka 0,80, sedangkan KNN memberikan hasil terendah namun cukup stabil pada kisaran 0,60-0,61. Temuan ini menegaskan bahwa Logistic Regression merupakan model tunggal yang paling efektif dalam memetakan sentimen publik untuk mencapai performa optimal.

Tabel 4. Hasil Evaluasi Dengan Tiga Pembagian Data

| Model               | Rasio | Accuracy | Precision | Recall | F1-Score |
|---------------------|-------|----------|-----------|--------|----------|
| KNN                 | 70:30 | 0.61     | 0.71      | 0.61   | 0.54     |
| Naive Bayes         | 70:30 | 0.80     | 0.80      | 0.80   | 0.80     |
| Logistic Regression | 70:30 | 0.84     | 0.84      | 0.84   | 0.84     |
| KNN                 | 65:35 | 0.60     | 0.69      | 0.60   | 0.54     |
| Naive Bayes         | 65:35 | 0.80     | 0.80      | 0.80   | 0.80     |
| Logistic Regression | 65:35 | 0.84     | 0.84      | 0.84   | 0.84     |
| KNN                 | 60:40 | 0.60     | 0.70      | 0.60   | 0.53     |
| Naive Bayes         | 60:40 | 0.80     | 0.80      | 0.80   | 0.80     |
| Logistic Regression | 60:40 | 0.84     | 0.84      | 0.84   | 0.84     |

#### 4.5. Hasil Penerapan Ensemble

Bagian ini menyajikan hasil klasifikasi sentimen menggunakan metode Ensemble Majority Voting yang menggabungkan prediksi dari tiga algoritma dasar: K-Nearest Neighbor, Naïve Bayes, dan Logistic Regression. Penggabungan ini dilakukan

menggunakan teknik Hard Voting untuk menentukan keputusan akhir berdasarkan suara terbanyak dari ketiga model. Hasil pengujian model *ensemble* pada berbagai skenario rasio data disajikan pada Tabel 5 berikut:



Tabel 5. Hasil Evaluasi Ensemble

| Ensemble (Majority Voting) |          |           |        |          |
|----------------------------|----------|-----------|--------|----------|
| Rasio                      | Accuracy | Precision | Recall | F1-Score |
| 70:30                      | 0.82     | 0.82      | 0.82   | 0.81     |
| 65:35                      | 0.82     | 0.82      | 0.82   | 0.82     |
| 60:40                      | 0.82     | 0.82      | 0.82   | 0.81     |

Berdasarkan data pada Tabel 5, model ensemble menunjukkan performa yang sangat stabil dengan akurasi sebesar 0,82 secara konsisten di seluruh skenario pengujian. Meskipun jumlah data latih dikurangi hingga rasio 60:40, nilai F1-Score tetap terjaga pada angka 0,81, yang mengindikasikan bahwa model penggabungan ini memiliki ketahanan yang baik terhadap variasi ukuran dataset.

#### 4.6. Analisis Perbandingan Seluruh Model

Evaluasi menunjukkan Logistic Regression sebagai model tunggal terunggul dengan akurasi 0,84, sedangkan KNN mencapai hasil terendah (0,61) akibat dimensi fitur teks yang tinggi. Perbandingan performa seluruh model pada rasio 70:30 disajikan pada Tabel 6 berikut ini:

Tabel 6. Komparasi Metrik Evaluasi (Rasio 70:30)

| Model Klasifikasi   | Accuracy | Precision | Recall |
|---------------------|----------|-----------|--------|
| KNN                 | 0.61     | 0.71      | 0.61   |
| Naïve Bayes         | 0.80     | 0.80      | 0.80   |
| Logistic Regression | 0.84     | 0.84      | 0.84   |
| Ensemble (Voting)   | 0.82     | 0.82      | 0.82   |

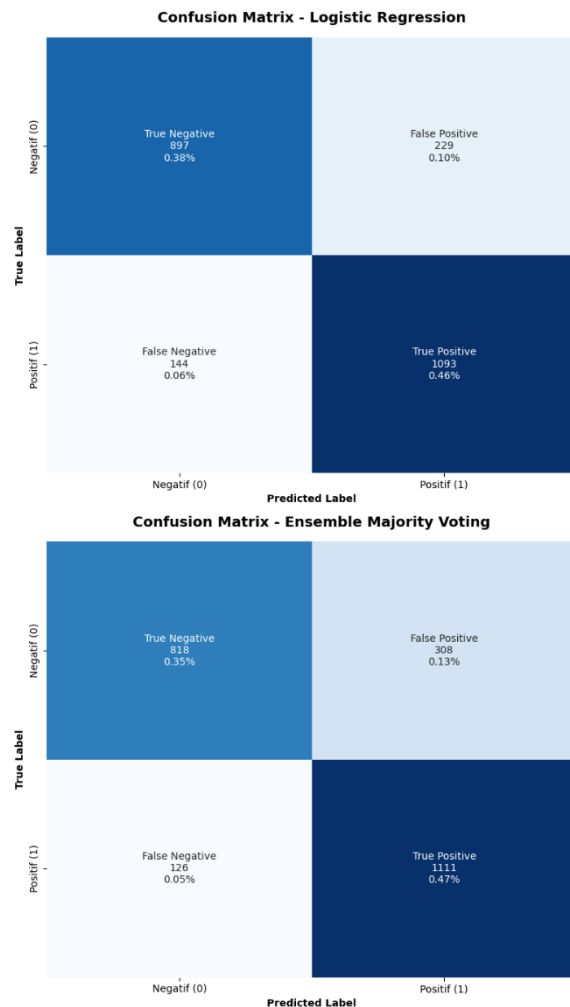
Berdasarkan pada Tabel 6, penerapan Ensemble Majority Voting menghasilkan akurasi stabil sebesar 0,82 dan terbukti efektif memitigasi kelemahan KNN melalui mekanisme hard voting. Model memiliki sensitivitas yang lebih tinggi pada kelas positif dengan nilai Recall 0,87 dibandingkan kelas negatif sebesar 0,80. Hal ini dipengaruhi oleh kebijakan pelabelan di mana skor netral digabungkan ke dalam kategori positif. Secara keseluruhan, metode ensemble terbukti lebih stabil dan kredibel dalam memetakan dinamika opini publik terkait kebijakan Dedi Mulyadi.

#### 4.7. Analisis Confusion Matrix Rasio Terbaik

Analisis mendalam dilakukan pada rasio 70:30 sebagai skenario dengan performa optimal. Berdasarkan confusion matrix pada Gambar 5, Logistic Regression menunjukkan akurasi yang lebih tajam sebesar 0,84 dibandingkan model Ensemble Majority Voting yang mencapai 0,82. Secara spesifik, Logistic Regression hanya mencatatkan 144 false negative dan 229 false positive, sedangkan model Ensemble mencatatkan 126 false negative dan 308 false positive.

Hal ini membuktikan bahwa model linier memiliki sensitivitas yang lebih superior dalam mengklasifikasikan aspirasi publik terkait kebijakan Dedi Mulyadi. Performa ini juga didukung oleh kebijakan pelabelan di mana komentar netral dikategorikan sebagai sentimen positif untuk

memperkaya pola data pendukung guna meningkatkan kemampuan generalisasi model.



Gambar 5. Confusion matrix Logistic Regression dan Ensemble Majority Voting pada rasio 70:30

## 5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa opini publik pada komentar YouTube terhadap kebijakan Dedi Mulyadi didominasi oleh sentimen positif sebesar 52,4% (4.123 ulasan) dibandingkan sentimen negatif sebesar 47,6% (3.752 ulasan), yang mengindikasikan bahwa mayoritas masyarakat memberikan respons mendukung terhadap kebijakan pengiriman anak bermasalah ke barak militer tersebut. Berdasarkan evaluasi model klasifikasi, Logistic Regression terbukti menjadi algoritma tunggal yang paling optimal dengan akurasi konsisten di angka 84%, sementara Naive Bayes mencapai 80% dan KNN berada pada kisaran 60%–61%. Penerapan Ensemble Majority Voting memberikan performa yang sangat stabil dengan rentang akurasi 82% serta berhasil menangani kelemahan KNN melalui mekanisme hard voting, sehingga menghasilkan klasifikasi yang lebih tangguh dan kredibel dalam menghadapi variasi ukuran dataset. Secara keseluruhan, temuan ini mencatat efektivitas metode Ensemble dan Logistic

Regression dalam memetakan dinamika sentimen publik, yang diharapkan dapat menjadi dasar evaluasi objektif bagi pemerintah dalam penyusunan kebijakan publik di masa mendatang.

#### DAFTAR PUSTAKA

- [1] A. Mahesha, D. Anggraeni, and M. I. Adriansyah, "Mengungkap Kenakalan Remaja: Penyebab, Dampak, dan Solusi," *Prim. J. Ilm. Multidisiplin*, vol. 2, no. 1, pp. 16–26, 2024, doi: 10.55681/primer.v2i1.278.
- [2] F. Afrita and F. Yusri, "Faktor-Faktor Yang Mempengaruhi Kenakalan Remaja," *Educ. J. Pendidik.*, vol. 2, no. 1, pp. 14–26, Dec. 2022, doi: 10.56248/educativo.v2i1.101.
- [3] Musta'a and R. Mesra, "Faktor Penyebab Kenakalan Remaja dan Cara Mengatasinya di Kelurahan Mentaya Seberang Seranau Kotawaringin Timur Kalimantan Tengah," *COMTE J. Sociol. Res. Educ.*, vol. 1, no. 5, pp. 239–249, 2024.
- [4] F. Bobyanti, "Kenakalan Remaja," *JERUMI J. Educ. Relig. Humanit. Multidiciplinary*, vol. 1, no. 2, pp. 476–481, 2023, doi: 10.57235/jerumi.v1i2.1402.
- [5] N. Elfemi, Y. Yuhelna, D. K. Anggreta, I. Isnaini, E. Erningsih, and S. Sarbaitinil, "Sosialisasi Penanggulangan Kenakalan Remaja: Upaya Preventif pada Remaja Awal," *INTAN CENDEKIA J. Pengabd. Masy.*, vol. 4, no. 1, pp. 30–36, 2023, doi: 10.47165/intancendekia.v4i1.638.
- [6] S. Azzahra, Z. Safardi, R. Adawiyyah, and B. Militer, "Membangun karakter disiplin melalui barak militer: analisis stakeholder dalam inovasi pendidikan," vol. 9, pp. 105–118, 2025.
- [7] D. Marganingsih, H. Oktavianto, and G. Abdurrahman, "Analisis Sentimen Komentar Youtube Masterchef Indonesia Menggunakan Algoritma Support Vector Machine dan Gaussian Naïve Bayes," *J. Inform. dan Teknol. Pendidik.*, vol. 5, no. 1, pp. 16–26, 2025, doi: 10.59395/jitp.v5i1.117.
- [8] M. P. Munthe, A. S. R. Ansori, and ..., "Analisis Sentimen Komentar Pada Saluran Youtube Food Vlogger Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes," *eProceedings Eng.*, vol. 8, no. 6, pp. 11909–11916, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/16897>
- [9] F. R. Dewi, "Klasifikasi Pola Peminjam Buku Berdasarkan Profesi Menggunakan Algoritma Naïve Bayes," vol. 15, no. 2, pp. 278–291, 2025.
- [10] M. Ujaran *et al.*, "Implementasi Convolutional Neural Network ( C NN ) Untuk," vol. 14, no. 2, pp. 314–325, 2024.
- [11] N. A. Laia and S. P. Barus, "Analisis Sentimen Pengguna Youtube Pada Video Berjudul '10 Tahun Jokowi Jadi Presiden,'" *JIKA (Jurnal Inform.)*, vol. 9, no. 2, p. 169, 2025, doi: 10.31000/jika.v9i2.13470.
- [12] M. Rizky Herdiansyah and A. Yuliana, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naïve Bayes Berdasarkan Komentar Pada Youtube," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 6, pp. 12454–12459, 2024, doi: 10.36040/jati.v8i6.11963.
- [13] D. N. Larasakti, A. Aziz, and D. Aditya, "Analisis Sentimen Komentar Video Youtube Dengan Metode K-Nearest Neighbor," *J. Ilm. Wahana Pendidik.*, vol. 2023, no. 5, pp. 132–142, 2023, [Online]. Available: <https://doi.org/10.5281/zenodo.7728573>
- [14] H. S. Mulyono and U. Saprudin, "Efektivitas Logistic Regression dalam Analisis Sentimen Berbahasa Indonesia pada Komentar YouTube tentang Isu," vol. 6, no. 3, pp. 1547–1555, 2025.
- [15] M. F. Fachrudin, C. V. Angkoso, and D. A. Fatah, "ANALISIS SENTIMEN PADA SOSIAL MEDIA TWITTER TERHADAP KUALITAS JARINGAN INTERNET TELKOMSEL MENGGUNAKAN ENSEMBLE K-NEAREST NEIGHBOUR -SUPPORT VECTOR MACHINE SENTIMENT ANALYSIS ON TWITTER SOCIAL MEDIA ON TELKOMSEL ' S INTERNET NETWORK QUALITY USING ENSEMBLE K-NEAREST NEIGHBOR -," vol. 11, no. 6, pp. 1253–1264, 2024, doi: 10.25126/jtiik.2024118713.
- [16] M. I. Ahmadi, D. Gustian, and F. Sembiring, "Analisis Sentiment Masyarakat terhadap Kasus Covid-19 pada Media Sosial Youtube dengan Metode Naive bayes," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 807–814, 2021.
- [17] F. Ahluna, C. J. Tutuarima, and I. Santoso, "Metode K-Nearest Neighbor Untuk Analisis Sentimen Tentang Penghapusan Ujian Nasional," *J. Ikraith-Informatika*, vol. 7, no. 2, pp. 1–6, 2023.
- [18] E. G. Radjah and A. C. Talakua, "Analisis Sentimen Komentar Terhadap Konten Tenun NTT di Youtube Menggunakan Metode SMOTE dan Logistic Regression," vol. XIII, no. November, pp. 84–94, 2024.
- [19] K. Nandini and M. Rahardi, "Sentiment Analysis of Economic Policy Comments on YouTube Using Ensemble Machine Learning," vol. 9, no. 5, 2025.
- [20] Z. Y. Burnama, M. A. Rosid, and N. L. Azizah, "Analisis Sentimen Pada Komentar Youtube Dalam Turnamen MPL Season 13 Dengan Metode Ensemble Machine Learning Sentiment Analysis on YouTube Comments in MPL Season 13 Tournament Using Ensemble Machine Learning Method".