

## ANALISIS SENTIMEN PUTUSAN MK SENGKETA PILPRES 2024 PADA YOUTUBE BERBASIS WEB DENGAN NAÏVE BAYES

Fakhri Setiawan<sup>1\*</sup>, Pipin Farida Ariyani<sup>2</sup>

<sup>1,2</sup>Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: <sup>1\*</sup>fakhrisetiawan128@gmail.com, <sup>2</sup>pipin.faridaariyani@budiluhur.ac.id  
(\* : *corresponding author*)

**Abstrak**-Kemajuan teknologi informasi telah memberikan dampak yang besar dalam mengubah media sosial menjadi platform utama bagi masyarakat untuk berbagi informasi, pandangan, dan opini mereka. Beragam platform media sosial yang ada seperti YouTube yang menjadi salah satu yang populer di Indonesia. Baru-baru ini, YouTube Indonesia sempat disibukkan karena pada tahun 2024, Indonesia menyelenggarakan Pemilihan Presiden dan Wakil Presiden, sebuah pesta demokrasi yang dilaksanakan setiap lima tahun sekali. Namun, seperti halnya dalam demokrasi di negara mana pun, pemilihan ini tidak lepas dari sengketa dan kontroversi dikarenakan putusan Mahkamah Konstitusi (MK). Tanggapan publik terhadap hasil putusan tersebut juga beragam, sebagian mendukung, sebagian netral, dan sebagian lagi menunjukkan ketidakpuasan atau ketidaksetujuan. Dengan demikian, dibutuhkan penelitian yang bertujuan untuk menganalisis sentimen terhadap komentar-komentar di YouTube khususnya pada saluran KompasTV dan CNN Indonesia yang menayangkan putusan MK terkait gugatan sengketa pemilihan presiden (Pilpres) 2024 di Indonesia. Melalui penelitian ini, dengan menggunakan metode Naïve Bayes untuk menganalisis sentimen terhadap komentar tersebut, dapat diperoleh pemahaman yang lebih baik tentang respons/opini masyarakat terhadap keputusan politik. Hasil dari analisis sentimen dengan rasio data latih 10% dan data uji 90% menunjukkan bahwa 63,8% masyarakat sentimen positif terhadap hasil putusan tersebut, 13,3% bersifat netral, dan 22,8% mengekspresikan sentimen negatif. Selain itu, penggunaan Naive Bayes menunjukkan akurasi mencapai 65%, presisi mencapai 62%, recall mencapai 73%, dan *f1-score* mencapai 61%.

**Kata Kunci:** YouTube, Analisis Sentimen, Mahkamah Konstitusi, Naïve Bayes

### SENTIMENT ANALYSIS OF THE CONSTITUTIONAL COURT'S DECISION ON THE 2024 PRESIDENTIAL ELECTION DISPUTE ON WEB-BASED YOUTUBE WITH NAIVE BAYES

**Abstract**-Advances in information technology have had a major impact in turning social media into the main platform for people to share their information, views and opinions. Various social media platforms exist such as YouTube, which is one of the most popular in Indonesia. Recently, YouTube Indonesia was preoccupied because in 2024, Indonesia held the Presidential and Vice Presidential Elections, a democratic party held every five years. However, as is the case in any democracy, this election was not free from disputes and controversy due to the Constitutional Court's ruling. The public's response to the verdict was also mixed, with some supportive, some neutral, and some showing dissatisfaction or displeasure. Thus, there is a need for research that aims to analyze the sentiment of comments on YouTube, especially on KompasTV and CNN Indonesia channels that broadcast the Constitutional Court's decision regarding the 2024 presidential election dispute lawsuit in Indonesia. Through this research, by using the Naïve Bayes method to analyze the sentiment of these comments, a better understanding of the public's response/opinion to political decisions can be obtained. The results of the sentiment analysis with a training and test data ratio of 90%:10% showed that 63.8% of the public had a positive sentiment towards the verdict, 13.3% were neutral, and 22.8% were negative.

**Keywords:** YouTube, Sentiment Analysis, Constitutional Court, Naïve Bayes

## 1. PENDAHULUAN

Kemajuan teknologi informasi telah mengakibatkan transformasi signifikan dalam cara masyarakat menyebarkan informasi dan mengungkapkan pandangan melalui media sosial. Media sosial adalah fitur internet yang memungkinkan pengguna untuk saling berbagi informasi [1]. Di antara berbagai platform yang tersedia, YouTube telah menjadi salah satu yang paling populer di Indonesia. Pada awal tahun 2024, pengguna YouTube di Indonesia mencapai 139 juta orang, menjadikan Indonesia sebagai negara dengan jumlah pengguna YouTube terbesar keempat di dunia. Rata-rata waktu penggunaan YouTube di Indonesia tercatat mencapai 31 jam 28 menit per bulan [2].

Peningkatan jumlah pengguna dan konten di YouTube telah menimbulkan beragam opini dan komentar dari pengguna, yang berperan penting dalam membentuk persepsi publik terhadap video yang diunggah. Komentar-

komentar ini tidak hanya mencerminkan pandangan individu tetapi juga dapat mempengaruhi penerimaan masyarakat terhadap konten tersebut [3]. Isu politik, seperti putusan Mahkamah Konstitusi dalam sengketa Pilpres 2024, telah memicu beragam reaksi publik. Putusan tersebut, yang seharusnya menjadi penentu dalam legitimasi hasil pemilihan, justru memicu kontroversi dan mengundang beragam opini dari masyarakat. Tanggapan publik terhadap hasil putusan Mahkamah Konstitusi bervariasi, mulai dari dukungan, sikap netral, hingga ketidakpuasan.

Dengan banyaknya komentar yang ada, analisis sentimen menjadi penting untuk mengidentifikasi dan mengukur perasaan serta opini publik terhadap isu ini. Melakukan analisis sentimen secara manual tidak mungkin dilakukan karena banyaknya data komentar yang tersedia. Oleh karena itu, diperlukan pendekatan sistematis dan efisien untuk mengklasifikasikan sentimen dalam komentar-komentar tersebut. Analisis sentimen memungkinkan untuk mengklasifikasikan komentar menjadi positif, netral, atau negatif, sehingga memberikan gambaran yang lebih jelas tentang bagaimana masyarakat merespons keputusan Mahkamah Konstitusi [4].

Beberapa penelitian sebelumnya telah dilakukan dalam bidang analisis sentimen media sosial, khususnya pada data komentar *YouTube*. Penelitian selanjutnya melakukan analisis sentimen komentar *YouTube* pada program Kampus Merdeka menggunakan algoritma *multinomial naïve bayes*, yang menghasilkan akurasi sebesar 86,11% [5]. Penelitian lain meneliti analisis sentimen komentar *YouTube* terkait konferensi tingkat tinggi G20 menggunakan metode *naïve bayes*, dengan nilai akurasi 77%, *F1-score* 76%, presisi 85%, dan *recall* 69% [3]. Penelitian selanjutnya juga melakukan penelitian serupa pada komentar *video* tentang mobil listrik di *YouTube* dengan hasil akurasi 70,69%, presisi 43,64%, dan *recall* 39,48% [6]. Selain itu, penelitian lain melakukan survei tentang analisis sentimen di media sosial menggunakan berbagai metode seperti *naïve bayes classifier*, *support vector machine*, K-NN, RNN, C4.5, *lexicon based* dan *LDA based topic modeling*, yang menunjukkan bahwa algoritma *Naïve Bayes* mendapatkan hasil akurasi tinggi [7]. Penelitian selanjutnya dalam analisis sentimen ulasan pengguna aplikasi iPusnas di *Google Play Store* menggunakan *multinomial naïve bayes* dengan rasio data latih dan data uji 9:1, mendapatkan akurasi 96,3%, presisi 96,67%, dan *recall* 96,67% [1].

Berdasarkan penelitian-penelitian tersebut, metode *naïve bayes* dipilih untuk penelitian ini karena kemampuannya dalam mengklasifikasikan sentimen. Penelitian ini menggunakan rasio perbandingan data latih dan data uji 9:1 untuk mengklasifikasikan sentimen opini terkait putusan Mahkamah Konstitusi dalam sengketa Pilpres 2024 di media sosial *YouTube*. Penelitian yang dilakukan sebelumnya, memiliki perbedaan tahapan dalam *preprocessing*, khususnya pada tahap penggantian *slangwords* [3]. Pada penelitian sebelumnya, proses penggantian *slangwords* tidak dilakukan, sementara dalam penelitian ini, tahap tersebut ditambahkan dengan contoh proses penerapan dari kata "mk" menjadi kata "mahkamah konstitusi". Penelitian ini juga mengukur akurasi, presisi, *recall*, dan *F1-score* dari implementasi algoritma *naïve bayes*.

Penelitian ini bertujuan untuk menganalisis sentimen terhadap komentar di *YouTube* KompasTV dan CNN Indonesia yang terkait dengan putusan Mahkamah Konstitusi dalam sengketa Pilpres 2024 di Indonesia. Fokus utama adalah mengidentifikasi apakah komentar tersebut bersifat positif, netral, atau negatif serta menguji akurasi metode *naïve bayes* dalam klasifikasi sentimen. Dengan demikian, penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik tentang respons masyarakat terhadap keputusan politik tersebut dan berkontribusi pada pengembangan metode analisis sentimen di media sosial.

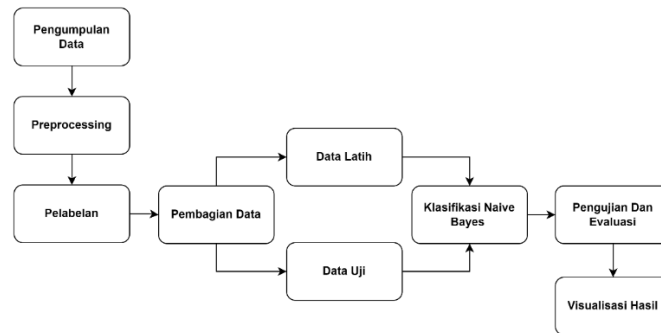
## 2. METODE PENELITIAN

### 2.1 Data Penelitian

*Dataset* dalam penelitian ini diambil dalam rentang waktu 22 April 2024 sampai 24 Mei 2024 dari komentar pada *video YouTube* di saluran KOMPASTV dan CNN Indonesia, yang berjudul "*BREAKING NEWS - Sidang Putusan MK Sengketa Pilpres 2024, Penentuan Anies Prabowo Ganjar!*" dan "*BREAKING NEWS! Sidang Putusan MK Terkait Sengketa Pilpres 2024.*" Data yang diambil meliputi "username", "created\_at", "text".

### 2.2 Penerapan Metode

Penelitian ini melibatkan beberapa tahapan dalam pembuatan sistem analisis sentimen dengan metode *naïve bayes*, yang mencakup seluruh proses dari awal hingga sistem berfungsi sepenuhnya, seperti yang ditunjukkan pada Gambar 1.



**Gambar 1.** Tahapan Metode

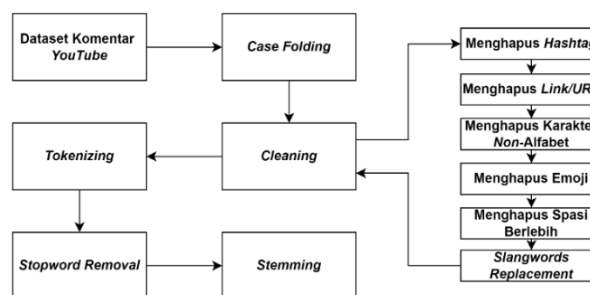
Penelitian ini dimulai dengan pengumpulan data komentar *YouTube* menggunakan teknik *crawling* di *Google Colab*, yang dilakukan terpisah dari program *website* utama. Data yang diperoleh disimpan dalam format *CSV* dan diintegrasikan ke dalam basis data. Selanjutnya, tahap *preprocessing* dilakukan untuk menyaring, mengeliminasi, dan memperbaiki kata-kata yang kurang relevan, menghasilkan teks bersih yang lebih terstruktur. Setelah itu, teks yang telah diproses dilabeli untuk menentukan sentimen positif, negatif, atau netral secara otomatis dengan membandingkan kamus kata dan memvalidasi secara manual. Setelah pelabelan, data dibagi menjadi dua kelompok utama, yaitu data uji dan data latih, sesuai dengan rasio yang ditentukan, menggunakan metode *stratified random sampling*. Proses selanjutnya adalah klasifikasi menggunakan metode *naïve bayes*. Tahap evaluasi kemudian dilakukan, yang mencakup pengujian dan analisis, dan hasilnya divisualisasikan dalam bentuk grafik untuk memberikan gambaran yang jelas.

### 2.3 Pengumpulan Data

Pengumpulan data dimulai dengan melakukan *crawling* dari *platform YouTube* melalui saluran *KOMPASTV* dan *CNN Indonesia*, memanfaatkan *YouTube data API v3*. Proses ini dimulai dengan memperoleh *API Key* melalui *Google Developers Console*, di mana *APIs* dan *Services* harus diaktifkan untuk memberikan akses ke *API Library*. *Crawling* dilakukan di luar program utama menggunakan *Google Colab*. Komentar *video YouTube* yang dikumpulkan kemudian disimpan dalam format *CSV (.csv)* dan diintegrasikan ke dalam basis data *SQL*.

### 2.4 Preprocessing

*Preprocessing* adalah langkah persiapan data yang sangat penting sebelum memasuki tahap pemodelan dan klasifikasi [8]. Proses ini mencakup pembersihan data dari kata-kata atau komentar *YouTube* yang tidak relevan dan tidak memiliki makna [9]. Ilustrasi proses tahapan *preprocessing* dalam penelitian ini dilihat pada Gambar 2.



**Gambar 2.** Tahapan *Preprocessing* Data

a. *Case Folding*

*Case Folding* adalah proses semua huruf dalam teks diubah menjadi huruf kecil [1].

b. *Cleaning*

*Cleaning* adalah proses pembersihan teks yang melibatkan penghapusan karakter yang tidak relevan, seperti tanda baca, simbol, atau URL [5]. Dalam tahap ini juga terdapat *slangword replacement* yang merupakan kata-kata berupa singkatan, bahasa gaul diubah menjadi kata yang benar.

c. *Tokenizing*

*Tokenizing* adalah tahap untuk memisahkan sebuah kalimat menjadi potongan kata per kata setelah melalui proses *case folding* [4]. Dengan kata lain, proses mengubah teks menjadi token [10].

#### d. Stopwords Removal

*Stopword removal* adalah proses menghapus kata-kata yang tidak relevan dengan sentimen dan yang tidak memiliki arti penting [8]. Untuk menghilangkan *stopwords* dapat menambahkan *library* [11].

#### e. Stemming

*Stemming* adalah proses mengubah kata menjadi bentuk kata dasar [12].

### 2.5 Pelabelan

Pada tahap pelabelan, komentar yang telah melewati *preprocessing* dilabeli secara otomatis dengan membandingkan kamus kata dan kemudian divalidasi secara manual. Terdapat tiga kategori label yang digunakan: positif, negatif, dan netral. Label positif diberikan kepada komentar yang menunjukkan dukungan atau kepuasan terhadap keputusan Mahkamah Konstitusi, sementara label negatif mencakup komentar yang menyatakan ketidakpuasan atau ketidaksetujuan. Label netral diterapkan pada komentar yang tidak menunjukkan dukungan atau ketidakpuasan yang jelas.

### 2.6 Pembagian Data

Pada pembagian data, komentar yang telah dilabeli dibagi menjadi 90% data latih dan 10% data uji menggunakan *stratified random sampling* yang merupakan teknik pembagian *dataset* yang didasarkan pada label tertentu. Data dipilih secara acak sederhana, sehingga setiap data memiliki kesempatan yang sama untuk terpilih dan untuk memastikan bahwa pembagian *dataset* dilakukan secara proporsional untuk setiap kelas yang ada [13]. Data latih digunakan untuk mengembangkan model dengan algoritma *naïve bayes*, sementara data uji digunakan untuk mengukur keakuratan prediksi model tersebut.

### 2.7 Klasifikasi Naïve Bayes

*Naïve bayes* adalah metode klasifikasi yang menggunakan prinsip probabilitas dan statistik bertujuan untuk memprediksi peluang di masa depan berdasarkan data dari pengalaman sebelumnya dengan asumsi dasar bahwa nilai-nilai atribut bersifat kondisional saling independen satu sama lain jika nilai *output* diberikan [14]. Bentuk umum dari klasifikasi *naïve bayes* dapat ditemukan dalam persamaan 1 berikut [13].

$$P(c|w) = \frac{P(w|c) \cdot P(c)}{P(w)} \quad (1)$$

Keterangan:

- $P(c|w)$  = *Posterior*, yaitu peluang suatu kelas  $c$  terhadap kata  $w$
- $P(w|c)$  = *Likelihood*, yaitu peluang kata  $w$  diberikan kelas  $c$
- $P(c)$  = *Prior*, yaitu peluang kemunculan kelas  $c$
- $P(w)$  = *Evidence*, yaitu peluang kemunculan kata  $w$

### 2.8 Rancangan Pengujian

Penelitian ini menguji model *naïve bayes* dengan menggunakan *confusion matrix* untuk menganalisis kinerja algoritma dalam tugas klasifikasi. Selain itu, *confusion matrix* juga merupakan metode yang umum digunakan untuk menghitung metrik seperti akurasi, *recall*, presisi, dan *f1-score*. Gambaran tentang tabel *confusion matrix* dapat dilihat pada tabel 1 yang disediakan [15].

**Tabel 1.** *Confusion Matrix*

<i>Confusion Matrix</i>	Prediksi			
	Positif	Negatif	Netral	
Aktual	Positif	TPos	FPosNeg	FPosNet
	Negatif	FNegPos	TNeg	FNegNeg
	Netral	FNetPos	FNegNeg	TNet

#### a. Akurasi

Akurasi adalah metrik evaluasi yang digunakan untuk mengukur seberapa sering model klasifikasi membuat prediksi yang benar. Perhitungan dengan menggunakan persamaan 2.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

b. Presisi

Presisi adalah metrik evaluasi yang mengukur ketepatan prediksi positif dari model, yaitu proporsi prediksi positif yang benar-benar tepat. Perhitungan dengan menggunakan persamaan 3.

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

c. Recall

Recall merupakan metrik yang mengukur sensitivitas atau kemampuan model untuk mendeteksi semua instance positif yang ada dalam data. Perhitungan dengan menggunakan persamaan 4.

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

d. F1-Score

F-1 Score ialah rata-rata harmonis dari presisi dan recall. F-1 Score memberikan gambaran seimbang antara presisi dan recall, dan sangat berguna ketika terdapat ketidakseimbangan kelas. Perhitungan dengan menggunakan persamaan 5.

$$F1 - \text{Score} = 2 \cdot \frac{\text{Presisi} \cdot \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\% \quad (5)$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Tahap Pengumpulan Data

Penelitian ini dimulai dengan mengumpulkan data secara terpisah dari program utama, menggunakan *Google Colab*. Data diambil dari komentar *YouTube* pada saluran KOMPASTV dan CNN Indonesia, yang menayangkan hasil putusan Mahkamah Konstitusi terkait sengketa Pilpres 2024. Total, terdapat 1060 komentar yang berhasil dikumpulkan dari kedua saluran tersebut dengan rentang waktu 22 April 2024 sampai 24 Mei 2024.

#### 3.2 Tahap Preprocessing

Tahap *preprocessing* dilakukan setelah *dataset* terkumpul, meliputi langkah-langkah: *case folding*, *cleaning*, *tokenizing*, *stopwords removal* dan *stemming*.

**Tabel 2.** Hasil Tahapan *Preprocessing*

Proses	Hasil <i>Preprocessing</i>
Teks Asli	Kyq main bola menang kalah itu biasa, ada yg bermain fair ada yg Enda, bgtu jg kehidupan, apalg politik, siapapun pemimpinannya kita hanya jd penonton koq da dpt apa" cian deih, kecuali setiap warga dibagi 20jt dr korupsi 271 T ☹️ itung aja untuk dpt berapa thn kita tuh ☹️ gubraq
<i>Case Folding</i>	kyq main bola menang kalah itu biasa, ada yg bermain fair ada yg enda, bgtu jg kehidupan, apalg politik, siapapun pemimpinannya kita hanya jd penonton koq da dpt apa" cian deih, kecuali setiap warga dibagi 20jt dr korupsi 271 t ???? itung aja untuk dpt berapa thn kita tuh ???? gubraq
<i>Cleaning</i>	kayak main bola menang kalah itu biasa ada yang bermain fair ada yang tidak begitu juga kehidupan apalagi politik siapapun pemimpinannya kita hanya jadi penonton mengapa tidak dapat apa kasian deh kecuali setiap warga dibagi juta dari korupsi triliun hitung saja untuk dapat berapa tahun kita tuh gubraq
<i>Tokenizing</i>	[kayak, main, bola, menang, kalah, itu, biasa, ada, yang, bermain, fair, ada, yang, tidak, begitu, juga, kehidupan, apalagi, politik, siapapun, pemimpinannya, kita, hanya, jadi, penonton, mengapa, tidak, dapat, apa, kasian, deh, kecuali, setiap, warga, dibagi, juta, dari, korupsi, triliun, hitung, saja, untuk, dapat, berapa, tahun, kita, tuh, gubraq]
<i>Stopwords Removal</i>	kayak main bola menang kalah bermain fair tidak kehidupan politik pemimpinannya penonton tidak kecuali warga dibagi juta korupsi triliun hitung untuk tahun
<i>Stemming</i>	kayak main bola menang kalah main fair tidak hidup politik pimpin tonton tidak kecuali warga bagi juta korupsi triliun hitung untuk tahun

#### 3.3 Tahap Pelabelan

Setelah *preprocessing*, jumlah data berkurang menjadi 1051, kemudian dilakukan pelabelan otomatis dengan pengecekan manual. *Dataset* diklasifikasikan ke dalam tiga kelas: positif untuk dukungan, negatif untuk penolakan atau kritik, dan netral untuk komentar tanpa opini yang jelas.

### 3.4 Tahap Pembagian Data

*Dataset* yang dilabeli dibagi menjadi data latih dan uji menggunakan *Stratified Random Sampling*. Rasio pembagian yang digunakan adalah 1:9 (10% data uji dan 90% data latih), menghasilkan 946 data latih dan 105 data uji.

### 3.5 Tahap Klasifikasi *Naïve Bayes*

Setelah pelabelan, klasifikasi dilakukan dengan algoritma *Naïve Bayes* rumus pada persamaan 1. Berikut ini adalah tahapan-tahapan dalam proses tersebut:

#### a. Persiapan

Pada tahap persiapan, digunakan sampel data Uji tercantum dalam Tabel 3. Data tersebut telah melewati proses *preprocessing* dan pelabelan.

**Tabel 3.** Sampel Data Uji

No	Clean Text	Sentiment
1	moga hakim mahkamah konstitusi sehat	Positif
2	bagus deh hasil putus mahkamah keluarga gemblung merdeka daerah negara satu republik indonesia arti pintar cerdas ambil	Positif
3	pasang calon untuk indonesia dukung untuk pribadi mahkamah konstitusi untuk bangsa indonesia ya allah ya rob beri baik untuk bangsa indonesia amiin alaaahuma amin	Positif
4	mahkamah konstitusi takut petik buah tunggu terima	Negatif
5	moga hakim mahkamah konstitusi taubat	Positif
6	kecewa sidang dagel sidang lelucon sidang main belaka mahkamah konstitusi lihat urus bangsa negara mahkamah konstitusi faham palu bodam emas pesan ibu mega putus mahkamah konstitusi hasil pilih presiden kecewa sungguh kecewa	Negatif
7	ternyata benar sidang ubah pangung sandiwara efek pasti pilih presiden curang	Negatif
8	muncul putus sidang mahkamah konstitusi	Netral
9	pilih umum sidang mahkamah konstitusi	Netral
10	alhamdulillah putus mahkamah konstitusi benar pilih rakyat	Positif

#### b. Perhitungan *Prior*

Setelah mendapatkan data Uji, langkah berikutnya adalah menghitung *prior* untuk setiap kelas.

$$P(C) = \frac{\text{Jumlah data berlabel pada suatu kelas}}{\text{total kelas}} \quad (6)$$

$$P(\text{positif}) = 5/10 = 0.5$$

$$P(\text{netral}) = 2/10 = 0.2$$

$$P(\text{negatif}) = 3/10 = 0.3$$

#### c. Perhitungan *Likelihood*

Setelah memperoleh *prior*, langkah selanjutnya adalah menghitung *likelihood* dengan tahapan perhitungan sebagai berikut.

$$P(w|c) = \frac{\text{jumlah kata pada suatu label} + 1}{\text{jumlah kata pada label} + \text{jumlah kata unik seluruh}} \quad (7)$$

$$\text{Jumlah kemunculan kata} = 121$$

$$\text{Jumlah kata unik seluruh label} = 67$$

$$\text{Kemunculan kata positif} = 58$$

$$\text{Kemunculan kata netral} = 10$$

$$\text{Kemunculan kata negatif} = 53$$

Mencari nilai *likelihood* kata data uji label Positif :

$$P(\text{alhamdulillah} | \text{Positif}) = (1 + 1) / (58 + 67) = 0.0134$$

$$P(\text{mahkamah} | \text{Positif}) = (5 + 1) / (58 + 67) = 0.0573$$

$$P(\text{putusan} | \text{Positif}) = (1 + 1) / (58 + 67) = 0.0134$$

$$P(\text{konstitusi} | \text{Positif}) = (5 + 1) / (58 + 67) = 0.0403$$

$$P(\text{benar} | \text{Positif}) = (2 + 1) / (58 + 67) = 0.0201$$

$$P(\text{pilih} | \text{Positif}) = (2 + 1) / (58 + 67) = 0.0201$$

$$P(\text{rakyat} | \text{Positif}) = (1 + 1) / (58 + 67) = 0.0134$$

Mencari nilai likelihood kata data uji label Negatif :

$$P(\text{mahkamah} | \text{Negatif}) = (4 + 1) / (53 + 67) = 0.0417$$

$$P(\text{alhamdulillah} | \text{Negatif}) = (0 + 1) / (53 + 67) = 0.0083$$

$$P(\text{putus} | \text{Negatif}) = (0 + 1) / (53 + 67) = 0.0083$$

$$P(\text{konstitusi} | \text{Negatif}) = (4 + 1) / (53 + 67) = 0.0417$$

$$P(\text{benar} | \text{Negatif}) = (1 + 1) / (53 + 67) = 0.0167$$

$$P(\text{pilih} | \text{Negatif}) = (1 + 1) / (53 + 67) = 0.0167$$

$$P(\text{rakyat} | \text{Negatif}) = (0 + 1) / (53 + 67) = 0.0083$$

Mencari nilai *likelihood* kata data uji label Netral :

$$P(\text{alhamdulillah} | \text{Netral}) = (0 + 1) / (10 + 67) = 0.0129$$

$$P(\text{putus} | \text{Netral}) = (0 + 1) / (10 + 67) = 0.0129$$

$$P(\text{mahkamah} | \text{Netral}) = (2 + 1) / (10 + 67) = 0.0389$$

$$P(\text{konstitusi} | \text{Netral}) = (2 + 1) / (10 + 67) = 0.0389$$

$$P(\text{benar} | \text{Netral}) = (0 + 1) / (10 + 67) = 0.0143$$

$$P(\text{pilih} | \text{Netral}) = (1 + 1) / (10 + 67) = 0.0259$$

$$P(\text{rakyat} | \text{Netral}) = (0 + 1) / (10 + 67) = 0.0129$$

#### d. Perhitungan *Posterior*

Setelah menghitung *prior* dan *likelihood*, langkah berikutnya adalah menghitung *posterior*. Karena komponen *evidence* konstan untuk semua kelas dan sering diabaikan dalam perhitungan *naïve bayes*, penyederhanaan ini tidak mempengaruhi hasil akhir. Proses perhitungannya dapat dilihat pada Tabel 4.

**Tabel 4.** Hasil Perhitungan Klasifikasi *Naïve Bayes*

Label	$\log(P(c)) + \sum \log(P(w c))$	Hasil
Positif	$\log(0.5) + \log(0.0134) + \log(0.0573) + \log(0.0134) + \log(0.0403) + \log(0.0201) + \log(0.0201) + \log(0.0134)$	-27.521
Netral	$\log(0.2) + \log(0.0129) + \log(0.0129) + \log(0.0389) + \log(0.0389) + \log(0.0129) + \log(0.0259) + \log(0.0129)$	-29.156
Negatif	$\log(0.3) + \log(0.0417) + \log(0.0083) + \log(0.0083) + \log(0.0417) + \log(0.0167) + \log(0.0167) + \log(0.0083)$	-30.2098

Hasil perhitungan klasifikasi komentar cenderung ke kelas positif. Dengan demikian, komentar tersebut diklasifikasikan sebagai positif, dan label yang diprediksi sesuai dengan label aktual.

### 3.6 Tahap Pengujian Metode

Tahap penghitungan akurasi dalam penelitian ini menggunakan *confusion matrix* untuk membandingkan nilai aktual dengan prediksi. Tabel 5 menunjukkan *confusion matrix* yang dihasilkan dengan rasio 10:90. Proses perhitungan didetailkan pada rasio 10:90.

**Tabel 5.** *Confusion Matrix* Rasio 10:90

N=105	Prediksi Positif	Prediksi Netral	Prediksi Negatif
Aktual Positif	38	13	16
Aktual Netral	0	13	1
Aktual Negatif	2	5	17
Jumlah	40	31	34

$$\text{Akurasi} = \frac{(38 + 13 + 17)}{(105)} = \frac{68}{105} \times 100\% = 0.647 = 65\%$$

$$\text{Presisi} = \frac{0.95 + 0.419 + 0.5}{3} = \frac{0.623}{3} = 0.623 \times 100\% = 62\%$$

$$\text{Recall} = \frac{0.567 + 0.928 + 0.708}{3} = \frac{2.204}{3} = 0.734 \times 100\% = 73\%$$

$$\text{F1-Score} = \frac{(0.710 + 0.526 + 0.586)}{3} = \frac{1.823}{3} = 0.607 \times 100\% = 61\%$$

### 3.7 Tahap Analisa Hasil Pengujian

Hasil Pengujian disajikan dalam bentuk tabel pada Tabel 6.

**Tabel 6.** Analisa Hasil Keseluruhan Pengujian

Rasio	Akurasi	Presisi	Recall	F1-Score	Hasil <i>Confusion Matrix</i>			Jumlah Data
					Prediksi Positif	Prediksi Netral	Prediksi Negatif	
10%:90%	65%	62%	73%	61%	38	13	17	105
20%:80%	52%	48%	55%	61%	60	18	31	210
30%:90%	48%	50%	60%	53%	38	13	17	316

Rasio pembagian data 10%:90% menghasilkan model klasifikasi terbaik dengan akurasi 65%, presisi 62%, *recall* 73%, dan *F1-score* 61%. Analisis menunjukkan mayoritas sentimen terhadap gugatan sengketa Pilpres 2024 bersifat positif (63,8%), dengan sentimen netral dan negatif masing-masing sebesar 13,3% dan 22,8%.

### 3.8 Halaman Beranda

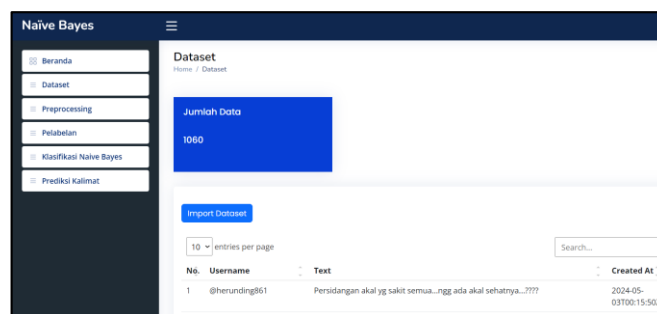
Antarmuka utama program menampilkan opsi menu pada *sidebar* untuk navigasi antara halaman aplikasi, serta informasi tentang jumlah *dataset* yang dimasukkan. Gambar 3 menggambarkan tampilan ini.



**Gambar 3.** Halaman Beranda

### 3.9 Halaman Dataset

Halaman tampilan *dataset* memungkinkan pengguna melihat dataset dalam bentuk mentah sebelum diproses oleh sistem. Pengguna dapat mengimpor dataset dengan tombol "*Import Dataset*" yang tersedia, dan file harus dalam format (.csv) agar berhasil diimpor. Terdapat juga fitur pencarian untuk memudahkan pengguna menemukan *dataset* tertentu. Gambar 4 menampilkan tampilan halaman ini.



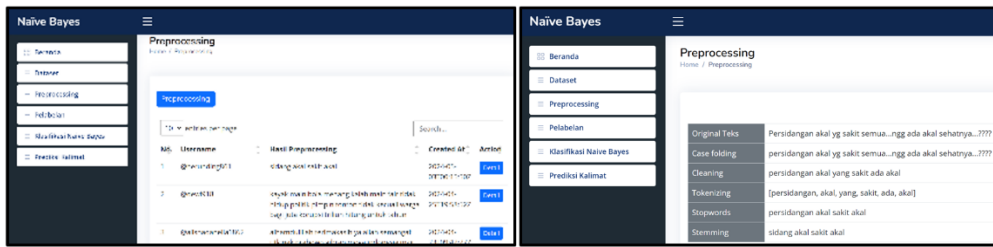
**Gambar 4.** Halaman Dataset

### 3.10 Halaman Preprocessing

Antarmuka ini digunakan untuk menjalankan proses pembersihan teks yang telah dilabeli secara otomatis. Pengguna dapat memulai proses ini dengan menekan tombol "*Preprocessing*". Selain itu, pengguna juga memiliki akses untuk memeriksa data setelah melalui proses *preprocessing*. Tahapan ini melibatkan proses *Case Folding*,



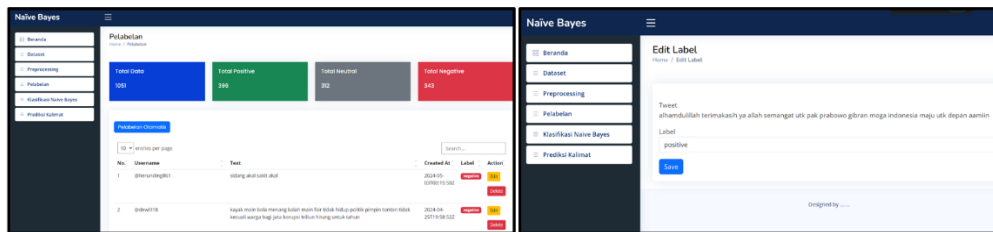
*Cleaning*, *Tokenizing*, penghapusan *Stopwords*, dan *Stemming*. Dengan mengklik tombol "Detail," pengguna dapat melihat setiap tahap dari proses *preprocessing* tersebut. Gambar 5 menggambarkan tampilan ini.



Gambar 5. Halaman *Preprocessing*

### 3.11 Halaman Pelabelan

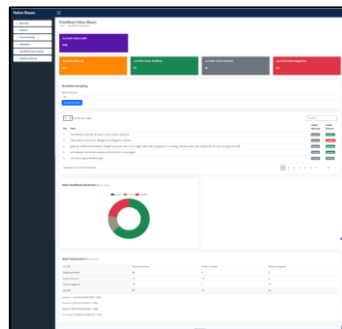
Gambar 6. menunjukkan detail fitur data pada halaman pelabelan, termasuk jumlah total setiap label dan data bersih setelah *preprocessing*. fitur *action* yang memungkinkan pengguna untuk mengganti, menambah, atau menghapus label, serta teks hasil *preprocessing* dengan *username*, tanggal pembuatan, dan label.



Gambar 6. Halaman Pelabelan

### 3.12 Halaman Klasifikasi *Naive Bayes*

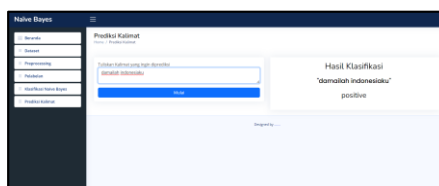
Pada antarmuka klasifikasi *Naive Bayes*, pengguna dapat melihat hasil klasifikasi setelah menekan tombol "Klasifikasi" dan mengatur rasio *sampling* sesuai keinginan. Antarmuka ini juga menampilkan grafik perbandingan label data dan perhitungan *confusion matrix*. Gambar 7 menampilkan tampilan halaman ini.



Gambar 7. Halaman Klasifikasi *Naive Bayes*

### 3.13 Halaman Prediksi Kalimat

Pada halaman prediksi kalimat, pengguna dapat memasukkan teks untuk diprediksi oleh sistem. Hasil prediksi muncul setelah pengguna menekan tombol "Mulai," yang menampilkan hasil sentimen dari teks tersebut. Gambar 8 menampilkan tampilan halaman ini.



Gambar 8. Halaman Prediksi Kalimat

#### 4. KESIMPULAN

Berdasarkan hasil pengujian dan evaluasi penelitian ini, dapat disimpulkan bahwa dari data komentar *YouTube* pada saluran KompasTV dan CNN Indonesia terkait putusan sengketa Pilpres 2024, sentimen masyarakat terbagi dalam tiga kategori yaitu, 63,8% menunjukkan sentimen positif, 13,3% bersifat netral, dan 22,8% mengekspresikan sentimen negatif. Hal ini menunjukkan bahwa mayoritas publik mendukung dan yakin bahwa putusan tersebut adil, meskipun ada sebagian yang bersikap netral dan tidak puas. Selain itu, metode *naive bayes* dengan rasio data latih 10% dan data uji 90% menghasilkan performa analisis sentimen terbaik, dengan akurasi 65%, presisi 62%, *recall* 73%, dan *F1-score* 61%. Hasil ini menunjukkan kemampuan *naive bayes* dalam mengklasifikasikan sentimen terhadap putusan Mahkamah Konstitusi, meskipun masih ada ruang untuk perbaikan lebih lanjut. Penelitian ini menyarankan penambahan *stopwords* dan *slangwords* sesuai konteks untuk meningkatkan kualitas *preprocessing* teks, serta pengembangan program fitur prediksi kalimat agar lebih efektif menangani teks di luar topik utama. Langkah-langkah ini diharapkan meningkatkan akurasi dan pemahaman terhadap sentimen masyarakat atas putusan Mahkamah Konstitusi terkait sengketa Pilpres 2024.

#### DAFTAR PUSTAKA

- [1] V. R. A. Sejati, et al, “Analisis Sentimen Ulasan Pengguna Aplikasi Ipusnas Pada Google Play Store Dengan Multinomial Naïve Bayes,” *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)* 21 Maret 2023, vol. 2, no. 1, 2023, , pp. 182-190.
- [2] We are social, “DIGITAL 2024,” we are social. [https://datareportal-com.translate.goog/reports/digital-2024-indonesia?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=id&\\_x\\_tr\\_hl=id&\\_x\\_tr\\_pto=tc](https://datareportal-com.translate.goog/reports/digital-2024-indonesia?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc)
- [3] R. A. Firsih, A. A. Arifiyanti, and D. S. Y. Kartika, “Analisis Sentimen Komentar Youtube Konferensi Tingkat Tinggi G20 Menggunakan Metode Naive Bayes,” *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 2, pp. 282–285, 2024, doi: 10.47233/jteksis.v6i2.1263.
- [4] C. A. Misrun, E. Haerani, M. Fikry, and E. Budianita, “Analisis Sentimen Komentar Youtube Terhadap Anies Baswedan Sebagai Bakal Calon Presiden 2024 Menggunakan Metode Naive Bayes Classifier,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 1, pp. 207–215, 2023, doi: 10.37859/coscitech.v4i1.4790.
- [5] R. A. Fauzan, and M. Mufti, “Analisis Sentimen Komentar Youtube Tentang Program Kampus Merdeka Berbasis Web Menggunakan Algoritma Multinomial Naïve Bayes,” *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, vol. 2, no. 2, September 2023, pp.864-871.
- [6] A. Karimah and G. Dwilestari, “Analisis Sentimen Komentar Video Mobil Listrik Di Platform Youtube Dengan Metode Naive Bayes,” *JATI (Jurnal Mahasiswa Teknik informatika)*, vol. 8, no. 1, pp. 767-773, 2024.
- [7] P. A. Permatasari, L. Linawati, and L. Jasa, “Survei Tentang Analisis Sentimen Pada Media Sosial,” *Majalah Ilmiah Teknologi Elektro*, vol. 20, no. 2, pp. 177-186, 2021, doi: 10.24843/mite.2021.v20i02.p01.
- [8] N. Ferdiansyah and A. Solichin, “Analisis Sentimen Terhadap Pembelajaran Dosen Berdasarkan Data Kritik Saran Mahasiswa Menggunakan Metode Naive Bayes,” *Bit (Fakultas Teknologi Informasi Universits Budi Luhur)*, vol. 19, no. 2, pp. 104-111, 2022.
- [9] A. A. Ningtyas, A. Solichin, and R. Pradana, “Analisis Sentimen Komentar Youtube Tentang Prediksi Resesi Ekonomi Tahun 2023 Menggunakan Algoritme Naïve Bayes,” *Bit (Fakultas Teknologi Informasi Universits Budi Luhur)*, vol. 20, no. 1, pp. 9-16, 2023.
- [10] M. I. Maulana, E. Budianita, M. Fikry, and F. Yanto, “Klasifikasi Sentiment Ulasan Aplikasi Sausage Man Menggunakan VADER Lexicon dan Naïve Bayes Classifier,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 3, pp. 485-492, 2023, doi: 10.30865/json.v4i3.5854.
- [11] R. Y. Lesmana, R. Andarsyah, “Model Klasifikasi Multinomial Naïve Bayes Untuk Analisis Sentiment Terkait Non-Fungible Token,” *Jurnal Teknik Informatika*, vol. 1, no. 3, pp. 135-139, 2022.
- [12] E. F. Saputra and M. R. Pribadi, “Analisis Sentimen Komentar Pada Kanal Youtube The Lazy Monday Menggunakan Algoritma Naïve Bayes” *MDP Student Conference (Msc) 2023 Universitas Multi Data Palembang*, pp. 17-23, 2023.
- [13] D. Juhaeni and A. Wibowo, “Penerapan Metode Naïve Bayes Untuk Wacana Kenaikan Harga Tiket Candi Borobudur Pada Twitter,” *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)* 6 September 2022, vol. 1, no. 1, 2022, pp. 683-690.
- [14] R. Safri Irawansyah, L. A. Syamsul Irfan A, and G. Wahyu Wiriasto, “Analisis Sentimen Terhadap Program Merdeka Belajar-Kampus Merdeka (Mbk) Pada Twitter Menggunakan Algoritma Naive Bayes Classifier (NBC),” *JTIKA: Jurnal Teknologi Informasi, Komputer dan Aplikasinya*, vol. 5, no. 2, pp. 237-244, 2023.
- [15] H. A. R. Harpizon et al., “Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes,” *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, pp. 131-140, 2022.