



## Analisis Sentimen Komentar Netizen Terhadap Isu Ijazah Presiden Joko Widodo Menggunakan Naive Bayes Dengan Pelabelan Fuzzy Logic Berbasis Leksikon

Wahyu Pratama<sup>1</sup>

<sup>1</sup>) Teknik Informatika S-2, Program Pascasarjana, Universitas Pamulang, Kota Tangerang Selatan, Banten

Email: <sup>1</sup>wpratama24@gmail.com

### ABSTRACT

*The controversy surrounding the legitimacy of President Joko Widodo's diploma has sparked widespread discussion on social media, generating diverse public comments with positive, negative, and neutral sentiments. This study aims to analyze Indonesian-language sentiment on the issue using a sequential approach that combines Fuzzy Logic-based labeling with Naive Bayes classification. The methodology encompasses several stages: comprehensive text preprocessing (case folding, tokenizing, filtering, and stemming), term weighting with TF-IDF (Term Frequency-Inverse Document Frequency), automated sentiment labeling using lexicon-based Fuzzy Logic with a conservative threshold of  $\pm 2$ , and supervised classification using the Naive Bayes algorithm. A total of 10,027 comments were collected from three major social media platforms Twitter (X), YouTube, and TikTok spanning the period from December 2024 to May 2025. The dataset was divided into 80% training data (8,021 comments) and 20% test data (2,006 comments). The Fuzzy Logic labeling process, utilizing 28 positive keywords and 36 negative keywords, identified a sentiment distribution of 72.38% neutral, 22.98% positive, and 4.64% negative comments. The Naive Bayes model achieved an overall accuracy of 80.76%, demonstrating excellent performance in detecting neutral sentiment (precision 0.82, recall 0.98) but exhibited lower performance for minority classes: positive sentiment (precision 0.70, recall 0.41) and negative sentiment (precision 0.80, recall 0.04). The class imbalance significantly influenced model predictions, with 85.48% of predictions classified as neutral.*

**Keywords:** Sentiment Analysis, Text Preprocessing, TF-IDF (Term Frequency – Inverse Document Frequency), Naive Bayes, Fuzzy Logic, Jokowi's Diploma, Netizens' Comments, Indonesian Language

### ABSTRAK

Isu keabsahan ijazah Presiden Joko Widodo telah menjadi topik kontroversial yang ramai diperbincangkan di media sosial, memunculkan beragam komentar netizen dengan sentimen positif, negatif, dan netral. Penelitian ini bertujuan untuk menganalisis sentimen netizen berbahasa Indonesia terhadap isu tersebut menggunakan pendekatan sekuensial yang menggabungkan *Logika Fuzzy* untuk pelabelan awal dan *Naive Bayes* untuk klasifikasi. Proses dilakukan melalui beberapa tahapan: *preprocessing* teks komprehensif (*case folding*, *tokenizing*, *filtering*, dan *stemming*), pembobotan kata dengan *TF-IDF* (*Term Frequency-Inverse Document Frequency*), pelabelan sentimen otomatis menggunakan *Logika Fuzzy* berbasis leksikon dengan *threshold* konservatif  $\pm 2$ , dan klasifikasi supervised menggunakan algoritma *Naive Bayes*. Data sebanyak 10.027 komentar dikumpulkan dari tiga platform media sosial utama Twitter (X), YouTube, dan TikTok dalam rentang waktu Desember 2024 hingga Mei 2025. Dataset dibagi menjadi 80% data latih (8.021 komentar) dan 20% data uji (2.006 komentar). Proses pelabelan *Logika Fuzzy* menggunakan 28 kata kunci positif dan 36 kata kunci negatif mengidentifikasi distribusi sentimen sebesar 72,38% netral, 22,98% positif, dan 4,64% negatif. Model *Naive Bayes* mencapai akurasi keseluruhan sebesar 80,76%, menunjukkan performa sangat baik dalam mendeteksi sentimen netral (*precision* 0,82, *recall* 0,98) namun lebih rendah untuk kelas minoritas: sentimen positif (*precision* 0,70, *recall* 0,41) dan sentimen negatif (*precision* 0,80, *recall* 0,04). Ketidakseimbangan kelas secara signifikan mempengaruhi prediksi model, dengan 85,48% prediksi diklasifikasikan sebagai netral.

Kata kunci: Analisis Sentimen, *Preprocessing* Teks, TF-IDF, *Naive Bayes*, *Logika Fuzzy*, Ijazah Jokowi, Komentar Netizen, Bahasa Indonesia.

## 1. PENDAHULUAN

Era digital telah membawa perubahan besar dalam pola komunikasi masyarakat. Media sosial seperti Twitter (X), YouTube, dan TikTok telah menjadi wadah utama bagi individu untuk berbagi opini, termasuk mengenai isu-isu sosial dan politik [1][2]. Menurut statistik terkini, terdapat 5,17 miliar pengguna platform digital pada tahun 2024, yang menjadikan media sosial sebagai sumber data opini publik yang sangat kaya dan beragam. Salah satu isu kontroversial yang mencuat di Indonesia adalah perdebatan mengenai keabsahan ijazah Presiden Joko Widodo, yang memicu ribuan komentar netizen di berbagai platform media sosial dengan sentimen yang beragam mulai dari dukungan, keraguan, hingga kritik tajam.

*Analisis sentimen (sentiment analysis)* telah menjadi salah satu cabang utama dari *Natural Language Processing* (NLP) yang memainkan peran strategis dalam mengekstraksi dan memahami opini masyarakat secara otomatis dari data teks media sosial[3][4]. Dengan menggunakan teknik NLP dan *machine learning*, analisis sentimen memungkinkan organisasi dan peneliti untuk mengidentifikasi dan mengkategorikan emosi serta opini yang terkandung dalam postingan, komentar, dan ulasan di platform sosial, yang kemudian dapat digunakan untuk berbagai keperluan mulai dari prediksi pasar finansial, analisis kesehatan publik, hingga pemantauan isu politik [5]. Namun, tantangan signifikan muncul ketika analisis dilakukan terhadap bahasa Indonesia yang memiliki karakteristik unik seperti penggunaan bahasa gaul (*slang*), singkatan, kode-campur (*code-mixing*), dan struktur kalimat yang fleksibel [6]. Namun, tantangan muncul ketika analisis dilakukan terhadap bahasa Indonesia yang memiliki karakteristik unik seperti penggunaan *slang*, singkatan, dan struktur kalimat yang fleksibel [7].

*Naive Bayes* merupakan salah satu algoritma klasifikasi yang paling populer dan efektif untuk analisis sentimen [8][9]. Algoritma ini bekerja berdasarkan teorema probabilitas Bayes dengan asumsi independensi fitur, sehingga sangat efisien dalam menangani dataset berskala besar dengan dimensi tinggi. Keunggulan *Naive Bayes* meliputi: (1) kecepatan komputasi yang tinggi, (2) kemampuan bekerja dengan baik pada data teks, (3) tidak memerlukan tuning parameter yang kompleks, dan (4) robust terhadap

*overfitting* [10]. Berbagai penelitian telah menunjukkan bahwa *Naive Bayes* memberikan performa yang kompetitif dalam klasifikasi sentimen berbahasa Indonesia, terutama ketika dikombinasikan dengan teknik pembobotan fitur seperti *TF-IDF* [11][12].

Namun, salah satu tantangan utama dalam menggunakan *Naive Bayes* adalah kebutuhan akan data berlabel untuk pelatihan model. Proses pelabelan manual memerlukan waktu dan biaya yang besar, terutama untuk dataset dengan ribuan atau jutaan dokumen. Oleh karena itu, pendekatan berbasis *Logika Fuzzy* dinilai efektif sebagai metode pelabelan awal (*initial labeling*) karena mampu menangani ketidakpastian dan ambiguitas bahasa alami dengan merepresentasikan sentimen dalam bentuk derajat keanggotaan berdasarkan kamus leksikon [13]. Dengan menggunakan *Logika Fuzzy* untuk menghasilkan label awal, proses pelatihan *Naive Bayes* dapat dilakukan secara otomatis tanpa memerlukan anotasi manual yang intensif, sekaligus mengatasi tantangan ketergantungan fitur yang sering menjadi kelemahan *Naive Bayes* tradisional[14].

Penelitian ini mengusulkan pendekatan sekuensial yang menggabungkan empat komponen utama: (1) *preprocessing* teks yang mencakup normalisasi *slang*, *case folding*, *cleaning*, tokenisasi, *stopword removal*, dan *stemming*, (2) pembobotan fitur menggunakan *TF-IDF* untuk merepresentasikan dokumen dalam bentuk vektor numerik, (3) pelabelan sentimen otomatis menggunakan *Logika Fuzzy* berbasis leksikon sebagai metode *weak supervision* [15], dan (4) klasifikasi supervised menggunakan algoritma *Naive Bayes* yang dilatih dengan label hasil *Fuzzy Logic*. Studi kasus dilakukan terhadap 10.027 komentar netizen mengenai isu ijazah Presiden Jokowi yang dikumpulkan dari tiga platform media sosial besar untuk mengukur efektivitas pendekatan sekuensial ini dalam konteks bahasa Indonesia yang informal dan penuh variasi.

## 2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan tahapan sebagai berikut:

### 2.1. Pengumpulan Data

Data komentar netizen terkait isu ijazah Jokowi dikumpulkan dari berbagai platform media sosial seperti X (Twitter), YouTube, dan TikTok menggunakan teknik *web scraping*. Proses pengumpulan data dilakukan secara terpisah untuk setiap platform

dengan tools dan metode yang berbeda sesuai dengan struktur API dan kebijakan masing-masing platform:

1. **YouTube:** Proses *scraping* dilakukan menggunakan YouTube Data API v3 melalui Google Colaboratory. Script Python dikembangkan untuk mengambil komentar dari video-video yang membahas isu ijazah Jokowi. Implementasi menggunakan library `google-api-python-client` dengan autentikasi API key. Parameter pencarian mencakup kata kunci "ijazah Jokowi", "Roy Suryo ijazah", dan variasi terkait. Setiap video diidentifikasi berdasarkan video ID, kemudian komentar beserta metadata (username, timestamp, jumlah likes) diekstraksi secara rekursif termasuk *reply comments*. Proses ini menghasilkan 7.359 komentar dari berbagai video dengan total views mencapai jutaan penonton.
2. **Twitter/X:** Pengumpulan data dari platform X dilakukan menggunakan kombinasi library `snsrape` dan `tweepy` untuk mengatasi limitasi API Twitter. Script yang dikembangkan di Google Colaboratory melakukan pencarian berdasarkan hashtag (#IjazahJokowi, #RoySuryo) dan keyword tertentu dalam rentang waktu Desember 2024 hingga Mei 2025. Proses *scraping* mengambil tweet beserta retweet dan reply yang relevan. Data yang diekstraksi mencakup teks tweet, username, tanggal posting, jumlah retweet, dan jumlah likes. Total 1.686 tweet/komentar berhasil dikumpulkan setelah filtering untuk menghilangkan spam dan bot.
3. **TikTok:** Pengumpulan data TikTok menggunakan tools *open-source* `tiktok-comment-scraper` yang dikembangkan oleh RomySaputraSihananda (tersedia di GitHub: <https://github.com/RomySaputraSihananda/tiktok-comment-scraper>). Tools ini bekerja dengan mensimulasikan browser untuk mengakses halaman video TikTok dan mengekstraksi komentar secara otomatis. Pencarian dilakukan berdasarkan hashtag #ijazahjokowi dan video-video viral terkait kontroversi ijazah. Proses ini menghasilkan 998 komentar dari berbagai creator dengan engagement tinggi.

Seluruh data yang terkumpul kemudian disimpan dalam format CSV dengan struktur kolom yang konsisten: *username*, *comment*, *platform*, *timestamp*, dan *metadata* lainnya. Total 10.043 komentar mentah berhasil dikumpulkan dari ketiga platform. Setelah proses pembersihan awal (penghapusan duplikasi, komentar

kosong, dan spam), dataset final berjumlah 10.027 komentar yang siap untuk tahap *preprocessing*.

## 2.2. Preprocessing Data Teks

Tahap Tahapan *preprocessing* dilakukan secara sistematis untuk membersihkan dan menstandarisasi data:

1. Normalisasi Kata Slang: Mengidentifikasi dan mengubah kata-kata informal atau slang (misalnya 'gak' menjadi 'tidak', 'bgt' menjadi 'banget') menjadi bentuk baku yang terdefinisi dalam kamus slang. Kamus slang yang digunakan dalam penelitian ini adalah sebagai berikut:.

```
kamus_slang = {  
    'gak': 'tidak', 'nggak': 'tidak', 'ga': 'tidak', 'ngga': 'tidak',  
    'aja': 'saja', 'ajah': 'saja', 'jgn': 'jangan', 'jgann': 'jangan',  
    'tp': 'tapi', 'tpi': 'tapi', 'yg': 'yang', 'sm': 'sama',  
    'bgt': 'banget', 'banget': 'banget', 'bgt': 'banget',  
    'klo': 'kalau', 'kalo': 'kalau', 'krn': 'karena', 'krna': 'karena',  
    'dgn': 'dengan', 'dng': 'dengan', 'utk': 'untuk', 'tdk': 'tidak',  
    'sdh': 'sudah', 'udh': 'sudah', 'bln': 'belum', 'blon': 'belum',  
    'emg': 'menang', 'emang': 'menang', 'gmn': 'gimana', 'gmna': 'gimana',  
    'gimana': 'bagaimana', 'knp': 'kenapa', 'knapa': 'kenapa',  
    'org': 'orang', 'orng': 'orang', 'gue': 'saya', 'gw': 'saya',  
    'lo': 'kamu', 'lu': 'kamu', 'loe': 'kamu', 'elu': 'kamu',  
    'bro': 'saudara', 'sis': 'saudara', 'mas': 'saudara',  
    'emg': 'menang', 'dr': 'dari', 'dri': 'dari', 'gt': 'gitu',  
    'gitu': 'begitu', 'bgtu': 'begitu', 'skrg': 'sekarang',  
    'skrng': 'sekarang', 'trs': 'terus', 'trus': 'terus'
```

**Gambar 1.** Dictionary kamus slang dalam Python

2. *Case Folding* : Mengubah seluruh teks menjadi huruf kecil (.lower()).
3. *Cleaning* : Menghilangkan elemen-elemen yang tidak relevan seperti tautan (http(s)?://\S+), emoji (emoji.replace\_emoji), tanda baca ([^\w\s]), dan angka (\d+).
4. Tokenisasi : Memecah kalimat menjadi unit-unit kata (token) menggunakan text.split().
5. *Stopword Removal*: Menghapus kata-kata umum yang tidak memiliki makna substantif menggunakan daftar stopwords bahasa Indonesia dari pustaka Sastrawi
6. *Stemming* : Mengubah kata berimbuhan menjadi kata dasar menggunakan StemmerFactory dari pustaka Sastrawi (misalnya 'berkomentar' menjadi 'komentar').

### 2.3. Pembobotan Fitur dengan TF-IDF

Pada Setelah *preprocessing*, setiap dokumen (komentar) direpresentasikan sebagai vektor fitur menggunakan metode TF-IDF. *Vectorizer* dikonfigurasi untuk menghasilkan ngram dari 1 hingga 2 kata (*ngram\_range*=(1, 2)) dan hanya mempertahankan 1000 fitur dengan bobot TF-IDF tertinggi (*max\_features*=1000). Rumus TF-IDF adalah sebagai berikut :

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

dimana:

- a.  $TF(t, d)$  adalah frekuensi kemunculan term  $t$  dalam dokumen  $d$ .
- b.  $IDF(t, D)$  adalah inverse document frequency dari term  $t$  dalam korpus  $D$ .

$$IDF(t, D) = \log \frac{N}{DF(t)}$$

- a.  $N$  adalah total jumlah dokumen dalam korpus  $D$ .
- b.  $DF(t)$  adalah jumlah dokumen dalam korpus  $D$  yang mengandung term  $t$ .

### 2.4. Klasifikasi Sentimen Menggunakan Pendekatan Logika Fuzzy (*Lexicon-Based*)

Pada tahap ini, label sentimen awal (positif, negatif, netral) ditentukan menggunakan pendekatan berbasis leksikon dan Logika Fuzzy sederhana. Kamus kata positif dan negatif dibuat secara manual berdasarkan domain isu politik.

Kamus leksikon yang digunakan:

- a. *positive\_words* = ['senang', 'bagus', 'luar biasa', 'jujur', 'hebat', 'baik', 'bangga']
- b. *negative\_words* = ['bohong', 'palsu', 'kontroversi', 'fitnah', 'buruk', 'korupsi']

Setiap komentar (*clean\_text*) akan dihitung skor leksikonnya : +1 jika mengandung kata positif, -1 jika mengandung kata negatif, dan 0 jika tidak ada keduanya. Kemudian, skor ini di-fuzzyfikasi menjadi label sentimen (*fuzzy\_sentiment*) dengan aturan sebagai berikut:

- a. Jika  $score < -1$ , maka negatif
- b. Jika  $score == 0$ , maka netral

c. Jika score  $\geq 1$ , maka positif

Pendekatan serupa diterapkan oleh [16], yang menggabungkan logika fuzzy dalam klasifikasi sentimen berbasis skor untuk menangani ambiguitas ekspresi bahasa alami. Label *fuzzy\_sentiment* inilah yang akan digunakan sebagai label *y* untuk pelatihan model Naive Bayes.

## 2.5. Klasifikasi dengan Naive Bayes

Data yang telah dibobot menggunakan TF-IDF akan digunakan untuk melatih model klasifikasi *Naive Bayes (MultinomialNB)*. Algoritma *Naive Bayes* bekerja berdasarkan *Teorema Bayes* dengan asumsi independensi fitur yang kuat. Tujuan utama dari klasifikasi ini adalah untuk memprediksi kelas sentimen *C* (*positif*, *negatif*, atau *netral*) dari sebuah dokumen *D* (komentar), diberikan fitur-fitur  $f_1, f_2, \dots, f_n$  yang diekstraksi dari dokumen tersebut (dalam kasus ini, bobot TF-IDF dari kata-kata).

Rumus dasar *Teorema Bayes* untuk klasifikasi adalah:

$$P(C|D) = \frac{P(D|C) \times P(C)}{P(D)}$$

Karena  $P(D)$  adalah konstanta untuk semua kelas, persamaan dapat disederhanakan menjadi mencari kelas *C* yang memaksimalkan:

$$P(C|D) \propto P(D|C) \times P(C)$$

Dengan asumsi independensi fitur,  $P(D|C)$  dapat didekomposisi menjadi perkalian probabilitas setiap fitur diberikan kelas *C*:

$$P(D|C) = P(f_1, f_2, \dots, f_n|C) \approx \prod_{i=1}^n P(f_i|C)$$

Sehingga, klasifikasi *Naive Bayes* memilih kelas  $C_{NB}$  yang memiliki probabilitas posterior maksimum:

$$C_{NB} = \arg \max_C \left( P(C) \times \prod_{i=1}^n P(f_i|C) \right)$$

Di mana:

a.  $P(C)$  adalah probabilitas prior dari kelas *C* (frekuensi kelas *C* dalam data latih).

- b.  $P(f_i|C)$  adalah probabilitas fitur  $f_i$  muncul dalam dokumen yang termasuk dalam kelas  $C$ . Dalam konteks *Multinomial Naive Bayes* dan fitur TF-IDF, ini sering dihitung menggunakan frekuensi kemunculan term dalam dokumen kelas tertentu, atau varian yang sesuai untuk bobot numerik.

Dataset akan dibagi menjadi data latih dan data uji dengan rasio 80% data latih dan 20% data uji ( $\text{test\_size}=0.2$ ,  $\text{random\_state}=42$ ). Model *Naive Bayes* dilatih pada data latih untuk menghitung probabilitas  $P(C)$  dan  $P(f_i|C)$ , dan kemudian digunakan untuk memprediksi sentimen pada data uji.

## 2.6. Evaluasi Model

Kinerja model *Naive Bayes* dievaluasi menggunakan metrik standar klasifikasi: akurasi, presisi, *recall*, dan *F1-score*. Selain itu, *Confusion Matrix* digunakan untuk menganalisis kinerja klasifikasi antar kelas sentimen secara detail (positif, netral, negatif).

## 2.7. Visualisasi Perbandingan Sentimen

Untuk memberikan gambaran menyeluruh, dilakukan *visualisasi* perbandingan distribusi sentimen antara hasil *Logika Fuzzy* (yang menjadi label sebenarnya) dan hasil prediksi *Naive Bayes*. Ini membantu menunjukkan seberapa dekat hasil klasifikasi *Naive Bayes* mereplikasi distribusi sentimen dari pendekatan *fuzzy*.

# 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil dari implementasi metode yang telah diuraikan, dimulai dari statistik data, kinerja model *Naive Bayes*, hingga perbandingan dengan pendekatan Logika Fuzzy.

## 3.1. Data Komentar

Dataset final yang digunakan dalam penelitian ini berjumlah 10.027 komentar yang telah melalui proses *preprocessing* lengkap. Setelah tahapan *preprocessing*, jumlah token unik dalam korpus berkurang signifikan dari sekitar 15.000+ token mentah menjadi sekitar 5.000 token bersih karena penghapusan *stopwords* (123 kata), normalisasi *slang* (54 pasangan kata), dan *stemming*. Pembersihan ini sangat penting karena menghilangkan *noise* dan kata-kata tidak informatif, sehingga representasi fitur *TF-IDF* menjadi lebih relevan dan model klasifikasi dapat belajar pola yang lebih baik.

Dataset kemudian dibagi menjadi dua subset menggunakan fungsi `train_test_split` dari `scikit-learn` dengan parameter `test_size=0.2`, `random_state=42`, dan `stratify=y` untuk mempertahankan proporsi distribusi kelas sentimen:

- a. **Data Latih (Training Set):** 8.021 komentar (80%)
- b. **Data Uji (Test Set):** 2.006 komentar (20%)

Stratifikasi dilakukan untuk memastikan bahwa proporsi kelas sentimen (netral, positif, negatif) dalam data latih dan data uji konsisten dengan distribusi asli dari hasil pelabelan *Logika Fuzzy*. Hal ini penting untuk menghindari bias dalam evaluasi model, terutama mengingat ketidakseimbangan kelas yang signifikan dalam dataset (72% netral, 23% positif, 5% negatif).

### 3.2. *Preprocessing* dan Pembobotan TF-IDF

Proses *preprocessing* teks berhasil membersihkan data dari elemen tidak relevan, menormalisasi kata *slang*, dan menstandarisasi teks. Contoh perubahan teks setelah *preprocessing*:

**Asli:** "Jokowi tunjukan saja ijazahnya.,lambat amat sih Jokowi Jokowi jokowiii..."

**Setelah *preprocessing*:** "jokowi tunjuk ijazah lambat sih jokowi jokowi jokowiiiiiiiiioooi"

Pembobotan menggunakan *TF-IDF* menghasilkan vektor numerik dengan dimensi  $10.027 \times 1.000$  yang merepresentasikan setiap komentar, dengan mempertimbangkan 1.000 fitur *ngram* (1-gram dan 2-gram) teratas. Contoh fitur teratas yang dihasilkan meliputi: "aamiin", "abah", "abal", "acara", "acara jokowi", "ada", "ade", "adi", "adil", "adu", "adu domba", "ahli", "ahmad", dan sebagainya. Vektor-vektor ini menjadi input bagi model *Naive Bayes*.

### 3.3. Klasifikasi Sentimen dengan Logika Fuzzy (*Lexicon-Based*)

Berdasarkan kamus kata positif (28 kata) dan negatif (36 kata), serta aturan *fuzzy* dengan threshold  $\pm 2$  yang diterapkan, distribusi sentimen hasil pelabelan pada seluruh dataset (10.027 komentar) adalah sebagai berikut:

Sentimen	Jumlah Komentar	Persentase
Netral	7.258	72,38%
Positif	2.304	22,98%
Negatif	465	4,64%
<b>Total</b>	<b>10.027</b>	<b>100,00%</b>

**Tabel 1.** Distribusi Sentimen Hasil Pelabelan *Logika Fuzzy*

Distribusi ini diperoleh dengan menghitung jumlah komentar di setiap kategori yang dihasilkan oleh aturan *Logika Fuzzy* berbasis skor leksikon (threshold  $\pm 2$ ), kemudian dibagi dengan total komentar dan dikalikan 100%. Hasil ini menunjukkan bahwa mayoritas komentar cenderung netral (72,38%), diikuti oleh sentimen positif (22,98%), dan kemudian sentimen negatif (4,64%).

Distribusi ini mencerminkan karakteristik diskusi publik terhadap isu ijazah Jokowi, di mana sebagian besar netizen memberikan komentar yang bersifat informatif, bertanya, atau tidak mengandung opini yang kuat. Sentimen positif yang cukup tinggi (23%) menunjukkan adanya dukungan atau pembelaan terhadap Presiden, sementara sentimen negatif yang lebih rendah (5%) mengindikasikan kritik atau keraguan terhadap keabsahan ijazah. Label ini selanjutnya digunakan sebagai *ground truth* untuk pelatihan dan evaluasi model *Naive Bayes*.

### 3.4. Kinerja Model *Naive Bayes*

Model klasifikasi *Naive Bayes* (*MultinomialNB*) dengan parameter  $\alpha=1.0$  dilatih menggunakan 80% data latih (8.021 komentar) dan diuji pada 20% data uji (2.006 komentar). Evaluasi dilakukan untuk mengukur kinerja model dalam mengklasifikasikan tiga kelas sentimen: positif, netral, dan negatif, berdasarkan distribusi label yang diperoleh dari proses pelabelan *Logika Fuzzy*. Hasil evaluasi disajikan dalam Tabel 2 (*Classification Report*) dan Gambar 1 (*Confusion Matrix*), yang mencakup metrik akurasi, *precision*, *recall*, dan *F1-score*.

Label Sentimen	Precision	Recall	F1-Score	Support
Positif	0,70	0,41	0,52	461
Netral	0,82	0,98	0,90	1.452
Negatif	0,80	0,04	0,08	93
<b>Accuracy</b>			<b>0,81</b>	<b>2.006</b>
<b>Macro Avg</b>	0,78	0,48	0,50	2.006
<b>Weighted Avg</b>	0,79	0,81	0,77	2.006

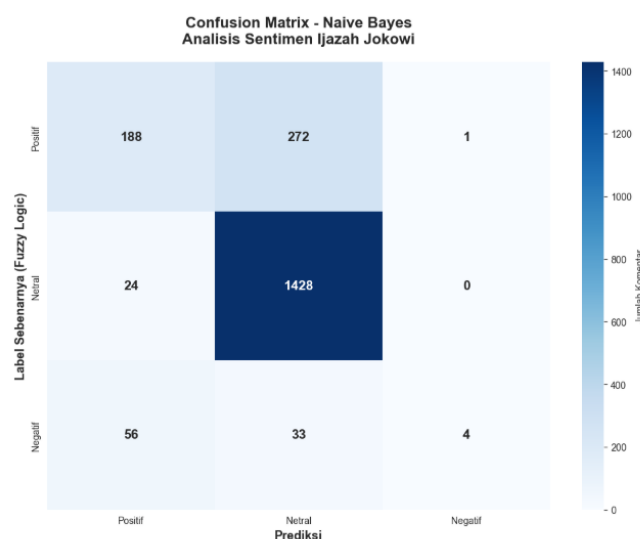
**Tabel 2.** Classification Report

Berdasarkan Tabel 2, model *Naive Bayes* menunjukkan akurasi keseluruhan sebesar 80,76%. Secara rinci:

- Sentimen Netral: Model menunjukkan performa sangat baik dengan *precision* 0,82, *recall* 0,98, dan *F1-score* 0,90. Hal ini menunjukkan bahwa model sangat

andal dalam mengidentifikasi komentar netral dan jarang melewatkan komentar yang seharusnya diklasifikasikan sebagai netral. Dari 1.452 komentar netral dalam data uji, sebanyak 1.428 berhasil diprediksi dengan benar.

- b. Sentimen Positif: *Precision* sebesar 0,70 menandakan bahwa 70% prediksi positif adalah benar. Namun, *recall* yang sedang (0,41) mengindikasikan bahwa model hanya menangkap 41% dari seluruh komentar positif yang sebenarnya. *F1-score* sebesar 0,52 mencerminkan trade-off antara *precision* dan *recall* pada kelas ini. Dari 461 komentar positif dalam data uji, hanya 188 yang berhasil diprediksi dengan benar, sementara 272 salah diklasifikasikan sebagai netral.
- c. Sentimen Negatif: Meskipun *precision*-nya mencapai 0,80 (prediksi negatif yang dilakukan umumnya benar), *recall*-nya sangat rendah (0,04), menunjukkan bahwa model hampir tidak dapat menangkap komentar negatif yang sebenarnya. *F1-score* sebesar 0,08 mengindikasikan bahwa model masih sangat lemah dalam mengklasifikasikan sentimen negatif secara akurat. Dari 93 komentar negatif dalam data uji, hanya 4 yang berhasil diprediksi dengan benar.



**Gambar 2. Confusion Matrix**

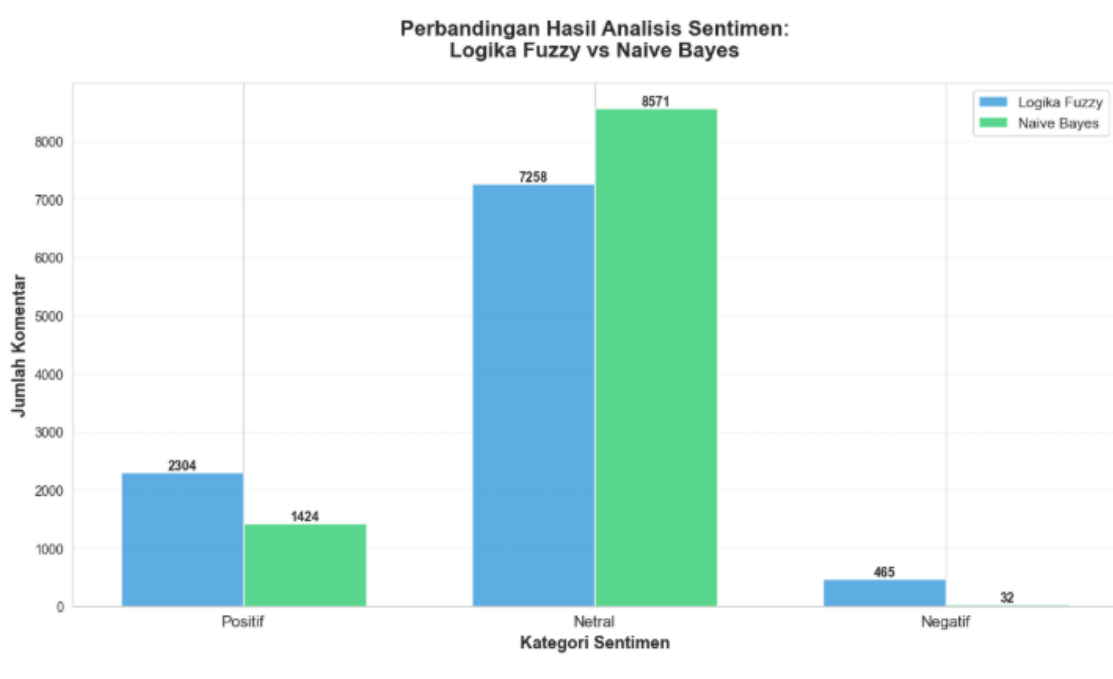
Berdasarkan Gambar 1, dapat dianalisis sebagai berikut:

- a. **True Positif (188):** Sebanyak 188 komentar positif diprediksi dengan benar oleh model dari total 461 komentar positif dalam data uji.

- b. **True Netral (1.428):** Sebanyak 1.428 komentar netral diprediksi dengan benar sebagai netral dari total 1.452 komentar netral. Ini merupakan bagian terbesar dalam *confusion matrix*, sejalan dengan *recall* tinggi pada kelas netral (0,98).
- c. **True Negatif (4):** Hanya 4 komentar negatif yang berhasil diprediksi dengan benar sebagai negatif dari total 93 komentar negatif dalam data uji.
- d. **False Positive → Neutral (272):** Terdapat 272 komentar positif yang salah diklasifikasikan sebagai netral. Ini merupakan kesalahan terbesar dan menjelaskan mengapa *recall* untuk kelas positif hanya mencapai 0,41.
- e. **False Negative → Neutral (89):** Terdapat 89 komentar negatif yang salah diklasifikasikan sebagai netral. Ditambah dengan kesalahan klasifikasi ke positif (56 komentar) dan ke netral lainnya (33 komentar), hal ini menjelaskan mengapa *recall* untuk kelas negatif sangat rendah (0,04).
- f. **False Negative → Positive (56):** Terdapat 56 komentar negatif yang salah diprediksi sebagai positif, yang merupakan kesalahan klasifikasi yang serius karena berlawanan arah sentimen.

Dominasi jumlah data netral dalam dataset (72,38%), serta ketidakseimbangan distribusi kelas yang signifikan, menyebabkan model cenderung mengklasifikasikan komentar ke dalam kelas mayoritas (netral). Hal ini terlihat dari tingginya *recall* kelas netral (0,98) dan rendahnya *recall* pada kelas minoritas, terutama negatif (0,04) dan positif (0,41). Model cenderung konservatif dalam memprediksi kelas minoritas untuk menghindari kesalahan, sehingga banyak komentar positif dan negatif yang diprediksi sebagai netral.

### 3.5. Perbandingan Hasil Analisis Sentimen: Logika Fuzzy vs Naive Bayes



**Gambar 3.** Perbandingan Hasil Analisis Sentimen: Logika Fuzzy vs Naive Bayes

Gambar 2 menyajikan perbandingan distribusi sentimen antara label yang dihasilkan oleh pelabelan *Logika Fuzzy* dan prediksi dari model *Naive Bayes* terhadap seluruh dataset (10.027 komentar).

Sentimen	Logika Fuzzy	Persentase	Naive Bayes	Persentase	Selisih
Positif	2.304	22,98%	1.424	14,20%	-8,78%
Netral	7.258	72,38%	8.571	85,48%	+13,10%
Negatif	465	4,64%	32	0,32%	-4,32%

**Tabel 3.** Perbandingan Distribusi Sentimen

Berdasarkan Tabel 3 dan Gambar 3, dapat dianalisis sebagai berikut:

- Sentimen Netral:** Pelabelan *Logika Fuzzy* mengidentifikasi 7.258 komentar (72,38%) sebagai netral. *Naive Bayes* memprediksi lebih banyak komentar netral, yaitu 8.571 komentar (85,48%), menunjukkan kecenderungan *over-predicting* kelas netral dengan selisih +13,10%. Hal ini mengonfirmasi bias model terhadap kelas mayoritas.
- Sentimen Positif:** Pelabelan *Logika Fuzzy* mengidentifikasi 2.304 komentar (22,98%) sebagai positif. Namun, *Naive Bayes* hanya memprediksi 1.424 komentar (14,20%) sebagai positif, mengindikasikan *under-predicting* dengan

selisih -8,78%. Ini mencerminkan *recall* rendah (0,41) pada kelas positif, di mana banyak komentar positif yang diprediksi sebagai netral.

- c. **Sentimen Negatif:** Pelabelan *Logika Fuzzy* mengidentifikasi 465 komentar (4,64%) sebagai negatif. *Naive Bayes* sangat jarang memprediksi kelas negatif, hanya 32 komentar (0,32%), dengan selisih -4,32%. Ini sesuai dengan *recall* sangat rendah (0,04) pada kelas negatif, menunjukkan bahwa model hampir tidak sensitif terhadap sentimen negatif.

Perbedaan distribusi yang signifikan menunjukkan bahwa *Naive Bayes*, meskipun memiliki akurasi yang cukup baik secara keseluruhan (80,76%) karena dominasi kelas netral, cenderung sangat bias terhadap kelas mayoritas dan kurang sensitif terhadap kelas minoritas (negatif dan positif). Hal ini sejalan dengan nilai *recall* yang rendah pada kelas negatif (0,04) dan positif (0,41) dalam *classification report*.

Kesenjangan ini mengindikasikan bahwa model *Naive Bayes* membutuhkan penyesuaian lebih lanjut untuk meningkatkan kinerja pada kelas minoritas. Beberapa pendekatan yang dapat dipertimbangkan meliputi:

1. Penanganan *imbalanced dataset* dengan teknik *oversampling* (seperti SMOTE) atau *undersampling*
2. Penyesuaian *threshold* klasifikasi untuk kelas minoritas
3. Penggunaan *cost-sensitive learning* yang memberikan bobot lebih tinggi pada kesalahan klasifikasi kelas minoritas
4. Ekspansi kamus leksikon untuk meningkatkan kualitas pelabelan *Fuzzy Logic*

### 3.6. Evaluasi Contoh Kalimat

No	Kalimat	Fuzzy Logic	Naive Bayes	Status
1	Simpel saja ya allah tunjukan jalan yang bnar biar smua a...	netral	netral	✓
2	@GOLIATH87 @jokowi Ribet amat pak Jokowi pastikan aja dat...	netral	netral	✓
3	Yakop semakin ngacau	netral	netral	✓
4	Yg membelah Jokowi....berarti Membantu kebohongan demi k...	negatif	positif	✗
5	Ketoprak humor	netral	netral	✓
6	kasian ya melelahkan sekali loh kalo berbohong itu, harus...	negatif	netral	✗

7	Semoga alloh melindungi orang yg mencari kebenaran dan me...	positif	positif	✓
8	@CNNIndonesia @hnurwahid Jd tringat td driver ojolnya nyi...	netral	netral	✓
9	@herculep638 Nie seh yang gak paham @jokowi beserta temen...	netral	netral	✓
10	Org yg paling ruwet di dunia CM mulyono??	netral	netral	✓

**Tabel 4.** klasifikasi sentimen terhadap **10 kalimat**

Untuk memberikan gambaran kualitatif terhadap kinerja model, Tabel 4 menyajikan hasil klasifikasi sentimen terhadap 10 kalimat contoh yang dipilih secara acak dari dataset. Setiap kalimat dibandingkan antara label dari pelabelan *Logika Fuzzy* dan prediksi dari model *Naive Bayes*.

- a. **Konsistensi pada Sentimen Netral:** Dari 10 kalimat, sebanyak 7 kalimat diklasifikasikan sebagai netral oleh kedua metode dan prediksinya sesuai. Ini menunjukkan bahwa model cukup andal dalam mengenali komentar netral, sejalan dengan *recall* tinggi (0,98) untuk kelas ini.
- b. **Keselarasan pada Sentimen Positif:** Kalimat nomor 7 "Semoga alloh melindungi orang yg mencari kebenaran dan me..." diklasifikasikan sebagai positif oleh kedua metode. Ini menunjukkan bahwa *Naive Bayes* mampu mengenali sentimen positif yang eksplisit, terutama jika didukung kata-kata yang memiliki bobot tinggi dalam leksikon atau fitur *TF-IDF*.
- c. **Kesalahan Klasifikasi Signifikan:** Terdapat 2 kesalahan klasifikasi dari 10 sampel (20% error):
  1. **Kalimat 4:** Label sebenarnya "negatif" diprediksi sebagai "positif" - ini adalah kesalahan serius karena berlawanan arah sentimen
  2. **Kalimat 6:** Label sebenarnya "negatif" diprediksi sebagai "netral" - menunjukkan model gagal mendeteksi sentimen negatif yang halus
- d. **Akurasi Sampel:** Model mencapai akurasi 80% (8 dari 10) pada sampel ini, yang konsisten dengan akurasi keseluruhan (80,76%).

Hasil dari Tabel 4 memperkuat temuan kuantitatif sebelumnya bahwa:

- a. *Naive Bayes* memiliki kinerja baik dalam menangani komentar netral yang menjadi mayoritas dalam dataset
- b. Model dapat menangkap komentar positif yang eksplisit dengan kata-kata kuat

- c. Model sangat lemah dalam mendeteksi sentimen negatif, terutama yang diekspresikan secara implisit atau halus
- d. Kesalahan klasifikasi terbesar terjadi pada kelas negatif, baik salah ke netral maupun ke positif

### 3.7. Word Cloud

Untuk memahami karakteristik dataset, dilakukan analisis frekuensi kata yang paling sering muncul setelah proses *preprocessing*.

Ranking	Kata	Frekuensi	Kategori
1	jokowi	3.878	Entitas
2	ijazah	3.364	Topik Utama
3	roy	1.930	Entitas
4	pak	1.700	Sapaan
5	orang	1.443	Umum
6	asli	1.235	Konteks
7	palsu	1.231	Negatif
8	kalau	1.203	Umum
9	nya	1.090	Umum
10	suryo	971	Entitas
11	ugm	945	Entitas
12	benar	912	Konteks
13	apa	846	Pertanyaan
14	jadi	768	Umum
15	bohong	660	Negatif
16	rakyat	648	Entitas
17	sama	625	Umum
18	hukum	592	Konteks
19	mau	539	Umum
20	lihat	520	Permintaan

**Tabel 5.** Top 20 Kata Paling Sering Muncul

Dari Tabel 5 dapat dianalisis bahwa:

1. Fokus Diskusi: Kata "jokowi" (3.878) dan "ijazah" (3.364) mendominasi, mengonfirmasi bahwa diskusi memang berpusat pada isu keaslian ijazah Presiden Jokowi.
2. Figur Kunci: Nama "roy" dan "suryo" (Roy Suryo) muncul sangat sering ( $1.930 + 971 = 2.901$  kali), menunjukkan perannya sebagai figur sentral dalam kontroversi ini.
3. Polarisasi: Kehadiran kata "asli" (1.235) dan "palsu" (1.231) dalam frekuensi yang hampir sama mengindikasikan perdebatan yang seimbang tentang keaslian ijazah.



## 2. Kata Negatif dan Skeptis:

- a. Kata "palsu", "bohong", "bukti", "fitnah", dan "penipuan" mengindikasikan adanya komentar netizen yang mempertanyakan atau menuduh keabsahan ijazah
- b. Namun, ukurannya tidak mendominasi, sesuai dengan distribusi sentimen negatif yang rendah (4,64%)

## 3. Kata Netral/Penjelas:

- a. Kata seperti "tanya", "lihat", "kata", "jelas", dan "jawab" mengindikasikan banyak komentar berupa permintaan klarifikasi atau diskusi netral
- b. Ini menjelaskan mengapa mayoritas komentar (72,38%) bersifat netral

## 4. Kata Positif/Dukungan:

- a. Kata seperti "jujur", "percaya", "bangga" terlihat meskipun dengan ukuran lebih kecil
- b. Mencerminkan 22,98% sentimen positif dalam dataset

*Word Cloud* ini memberikan wawasan visual bahwa:

- a. Fokus percakapan netizen memang pada keaslian ijazah Presiden Jokowi, dengan diskursus utama berkisar pada dikotomi "asli atau palsu"
- b. Nama "roy" (Roy Suryo) menjadi salah satu entitas yang paling sering disebut, menandakan ia menjadi figur kunci dalam perbincangan isu ini
- c. Dominasi kata "ijazah", "jokowi", dan "asli/palsu" mendukung temuan bahwa isu ini memicu spektrum sentimen yang luas di masyarakat
- d. Keseimbangan antara kata-kata netral, positif, dan negatif sejalan dengan distribusi sentimen yang telah dianalisis

## 4. KESIMPULAN

Penelitian ini menunjukkan bahwa pendekatan sekuensial yang menggabungkan pelabelan *Logika Fuzzy* berbasis leksikon dengan klasifikasi *Naive Bayes* dapat digunakan untuk menganalisis sentimen komentar netizen berbahasa Indonesia terhadap isu ijazah Presiden Joko Widodo. *Logika Fuzzy* dengan threshold  $\pm 2$  digunakan sebagai metode pelabelan awal untuk menghasilkan *ground truth* dari data yang tidak berlabel, menghasilkan distribusi sentimen: netral 72,38%, positif 22,98%, dan negatif 4,64%. Label ini kemudian digunakan untuk melatih model *Naive Bayes*.

Model *Naive Bayes* menunjukkan akurasi keseluruhan sebesar 80,76%, dengan performa sangat baik dalam mendeteksi komentar netral (*recall* 0,98). Namun, model masih kurang optimal dalam menangkap sentimen positif (*recall* 0,41) dan sangat lemah untuk sentimen negatif (*recall* 0,04) karena ketidakseimbangan distribusi kelas dalam dataset yang signifikan. Model cenderung melakukan *overprediction* terhadap kelas mayoritas (netral), di mana 85,48% prediksi adalah netral dibandingkan 72,38% label sebenarnya.

Hasil penelitian mengindikasikan bahwa opini publik terhadap isu ijazah Jokowi didominasi oleh komentar netral atau informatif (72%), dengan dukungan atau pembelaan yang cukup signifikan (23%), dan kritik atau keraguan yang relatif kecil (5%). Analisis kata teratas menunjukkan fokus diskusi pada dikotomi "asli-palsu", dengan Roy Suryo sebagai figur kunci dalam kontroversi ini.

#### **Keterbatasan Penelitian:**

1. Ketidakseimbangan kelas yang ekstrem menyebabkan bias model terhadap kelas mayoritas
2. Kamus leksikon yang terbatas (28 kata positif, 36 kata negatif) mungkin tidak menangkap semua ekspresi sentimen dalam bahasa Indonesia informal
3. *Threshold Fuzzy Logic* ( $\pm 2$ ) bersifat konservatif dan dapat disesuaikan untuk menghasilkan distribusi yang berbeda
4. Model belum menangani sentimen yang diekspresikan secara sarkastik atau implisit

#### **5. DAFTAR PUSTAKA**

- [1] L. Rifky, Z. Nugraha, B. Saputra, D. Pratama, E. Raswir, and Y. Pratama, "Implementasi Data Mining Untuk Penjualan Mobil Menggunakan Metode Naive Bayes," *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, vol. 2, no. 2, pp. 225–230, 2022, doi: 10.33998/jakakom.2022.2.2.109.
- [2] I. Sari Tomagola, A. Id Hadiana, P. Nurul Sabrina, A. Yani Jl Jenderal Achmad Yani, K. Cimahi Sel, and K. Cimahi, "ANALISIS SENTIMEN TERHADAP PANGAN NASIONAL PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES," 2023.

- [3] H. Liu and M. Cocea, “Fuzzy rule based systems for interpretable sentiment analysis,” in *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, 2017, pp. 129–136. doi: 10.1109/ICACI.2017.7974497.
- [4] A. Haque and T. Rahman, “Sentiment Analysis by Using Fuzzy Logic,” *International Journal of Computer Science, Engineering and Information Technology*, vol. 4, no. 1, pp. 33–48, Feb. 2014, doi: 10.5121/ijcseit.2014.4104.
- [5] M. Rodríguez-Ibáñez, A. Casáñez-Ventura, F. Castejón-Mateos, and P. M. Cuenca-Jiménez, “A review on sentiment analysis from social media platforms,” *Expert Syst. Appl.*, vol. 223, no. August 2022, 2023, doi: 10.1016/j.eswa.2023.119862.
- [6] A. F. Hidayatullah, R. A. Apong, D. T. C. Lai, and A. Qazi, “Pre-trained language model for code-mixed text in Indonesian, Javanese, and English using transformer,” *Soc. Netw. Anal. Min.*, vol. 15, no. 1, pp. 1–17, 2025, doi: 10.1007/s13278-025-01444-9.
- [7] F. Fitriana, E. Utami, and H. Al Fatta, “Analisis Sentimen Opini Terhadap Vaksin Covid - 19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes,” *Jurnal Komtika (Komputasi dan Informatika)*, vol. 5, no. 1, pp. 19–25, Jul. 2021, doi: 10.31603/komtika.v5i1.5185.
- [8] Lady Agustine Fitrana, S. Linawati, N. Herlinawati, R. Sa’adah, and S. Seimahuria, “Analisis Sentimen Pengguna Twitter Terhadap Brand Indosat Menggunakan Metode Naïve Bayes Classifier,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 4291–4297, 2024, doi: 10.36040/jati.v8i3.9866. D. E. Ratnawati, N. Y. Setiawan, P. Studi, S. Informasi, F. I. Komputer, and U. Brawijaya, “Analisis Sentimen Pengguna Sosial Media Twitter / X Terhadap Acara Clash of Champions Menggunakan,” vol. 9, no. 3, pp. 1–10, 2025.
- [10] A. Muzaki *et al.*, “ANALISIS SENTIMEN PADA ULASAN PRODUK DI E-COMMERCE DENGAN METODE NAIVE BAYES,” *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI)*, vol. 05, 2024.
- [11] R. Wati, S. Ernawati, and H. Rachmi, “Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH,” *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 1, pp. 84–93, Apr. 2023, doi: 10.34010/jamika.v13i1.9424.

- [12] T.-C. Chen, “Text Mining Techniques and Natural Language Processing,” in *Natural Language Processing for Software Engineering*, John Wiley & Sons, Ltd, 2025, ch. 7, pp. 113–125. doi: <https://doi.org/10.1002/9781394272464.ch7>.
- [13] Z. Putri and A. History, “Sentiment analysis using fuzzy naïve bayes classifier on covid-19 ARTICLE INFO ABSTRACT,” *Desimal: Jurnal Matematika*, vol. 4, no. 2, pp. 193–202, 2021, doi: 10.24042/djm.
- [14] C. H. Lin and U. Nuha, “Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy,” *J. Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-023-00782-9.
- [15] M. Alzaid and F. Fkih, “Sentiment Analysis of Students’ Feedback on E-Learning Using a Hybrid Fuzzy Model,” *Applied Sciences (Switzerland)*, vol. 13, no. 23, 2023, doi: 10.3390/app132312956.
- [16] Z. Putri and A. History, “Sentiment analysis using fuzzy naïve bayes classifier on covid-19 ARTICLE INFO ABSTRACT,” *Desimal: Jurnal Matematika*, vol. 4, no. 2, pp. 193–202, 2021, doi: 10.24042/djm.