

Analisis Sentimen Berbahasa Indonesia Menggunakan Preprocessing Teks, TF-IDF, Naive Bayes, dan Logika Fuzzy: Studi Kasus Komentar Netizen tentang Ijazah Jokowi

Wahyu Pratama^{*1},

¹Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang,
Tangerang Selatan, Indonesia
Email: [*1wpratama24@gmail.com](mailto:1wpratama24@gmail.com),

(Naskah masuk: 6 Agustus 2025, diterima untuk diterbitkan: 30 Januari 2026)

Abstrak: Isu keabsahan ijazah Presiden Joko Widodo telah menjadi topik kontroversial yang ramai diperbincangkan di media sosial, menimbulkan beragam komentar netizen dengan sentimen positif, negatif, dan netral. Penelitian ini bertujuan untuk menganalisis sentimen berbahasa Indonesia terhadap isu tersebut dengan menggunakan pendekatan hibrida. Metode yang digunakan mencakup preprocessing teks (case folding, tokenizing, filtering, dan stemming), pembobotan kata dengan TF-IDF, klasifikasi awal menggunakan algoritma Naive Bayes, serta penyempurnaan hasil klasifikasi melalui Logika Fuzzy untuk menangani ambiguitas dan ketidakpastian dalam bahasa alami. Data dikumpulkan dari Twitter, YouTube, dan TikTok, dengan total 10.248 komentar netizen, di mana 2.082 digunakan sebagai data uji. Hasil menunjukkan bahwa Naive Bayes mampu mengklasifikasikan sentimen dengan akurasi 82,1%, sementara integrasi Logika Fuzzy meningkatkan akurasi hingga 88,5%. Analisis distribusi sentimen akhir menunjukkan dominasi sentimen netral, disusul negatif, dan positif. Pendekatan ini terbukti efektif dalam menangani kompleksitas bahasa Indonesia dalam konteks opini publik, khususnya pada isu sosial-politik yang sensitif.

Kata Kunci – Analisis Sentimen, Preprocessing Teks, TF-IDF, Naive Bayes, Logika Fuzzy, Ijazah Jokowi, Komentar Netizen, Bahasa Indonesia

Abstract: The validity of President Joko Widodo's diploma has become a controversial topic that has been widely discussed on social media, causing various netizen comments with positive, negative, and neutral sentiments. The objective of this study is to analyze the sentiment in Indonesia regarding the issue by employing a hybrid approach. The methods employed encompass text preprocessing (case folding, tokenizing, filtering, and stemming), word weighting with TF-IDF, initial classification utilizing the Naive Bayes algorithm, and refinement of classification outcomes through Fuzzy Logic to address ambiguity and uncertainty in natural language. The data was collected from prominent social media platforms, including Twitter, YouTube, and TikTok. A total of 10,248 comments from netizens were gathered, of which 2,082 were utilized as test data for the study. The findings indicate that Naive Bayes demonstrates an accuracy of 82.1% in sentiment classification, while the incorporation of Fuzzy Logic enhances this accuracy to 88.5%. The final sentiment distribution analysis demonstrated the predominance of neutral sentiment, followed by negative and positive sentiments. This approach has proven effective in dealing with the complexity of the Indonesian language in the context of public opinion, especially on sensitive socio-political issues.

Keywords – Sentiment Analysis; Text Preprocessing; TF-IDF; Naive Bayes; Fuzzy Logic; Jokowi's Degree; Netizens' Comments; Indonesian Language (.)

1. PENDAHULUAN

Perkembangan media sosial telah mengubah cara masyarakat Indonesia menyampaikan opini terhadap isu-isu publik. Platform digital tidak hanya berfungsi sebagai sarana komunikasi, tetapi juga menjadi ruang terbuka bagi masyarakat untuk mengekspresikan sikap, kritik, dan dukungan terhadap kebijakan maupun figur publik. Salah satu isu yang memicu diskursus luas di ruang digital adalah polemik mengenai keabsahan ijazah Presiden Joko Widodo. Beragam komentar netizen yang muncul mencerminkan spektrum sentimen yang kompleks, mulai dari dukungan, keraguan, hingga penolakan yang disampaikan secara eksplisit.

Untuk memahami pola opini publik tersebut, analisis sentimen menjadi pendekatan yang banyak digunakan dalam bidang Natural Language Processing (NLP). Analisis sentimen bertujuan mengklasifikasikan opini ke dalam kategori tertentu, umumnya positif, negatif, dan netral, berdasarkan konten teks[1]. Pada

konteks bahasa Indonesia, proses ini menghadapi tantangan tersendiri akibat karakteristik bahasa yang dinamis, penggunaan slang, singkatan, serta ambiguitas makna yang sering muncul dalam komentar media sosial. Penelitian Kurniawan et al. (2020) menekankan bahwa tahap preprocessing teks yang komprehensif menjadi faktor krusial dalam meningkatkan kualitas analisis sentimen berbahasa Indonesia, mencakup proses case folding, tokenization, stopword removal, dan stemming [2].

Penelitian sebelumnya menunjukkan bahwa metode berbasis machine learning seperti Naive Bayes menjadi pilihan populer karena kesederhanaan dan efisiensinya dalam mengolah data teks berukuran besar [3]. Kombinasi Naive Bayes dengan pembobotan fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) terbukti mampu menghasilkan performa klasifikasi yang baik pada berbagai studi analisis sentimen berbahasa Indonesia [4]. Penelitian Khoerunnisa et al. (2025) menunjukkan bahwa kombinasi algoritma Naive Bayes dengan teknik TF-IDF dan Cross Validation mencapai akurasi optimal dalam mendeteksi sentimen positif pada data cuitan berbahasa Indonesia [5]. Sementara itu, studi Adharani et al. (2025) yang menganalisis sentimen terhadap Tim Nasional Indonesia berhasil mencapai akurasi 79,38% menggunakan Naive Bayes dengan SMOTE untuk menangani ketidakseimbangan data [6]. Namun, pendekatan klasifikasi konvensional ini cenderung kaku dan kurang mampu merepresentasikan sentimen yang ambigu atau berada di antara dua kelas.

Untuk mengatasi keterbatasan tersebut, beberapa penelitian mulai mengintegrasikan Logika Fuzzy ke dalam proses analisis sentimen. Logika Fuzzy memungkinkan representasi derajat keanggotaan suatu sentimen, sehingga opini tidak dipaksa masuk ke dalam kategori yang bersifat mutlak[7]. Pendekatan hibrida antara metode statistik dan logika fuzzy dilaporkan mampu meningkatkan ketepatan klasifikasi, terutama dalam menangani opini netral dan sentimen yang tidak tegas. Penelitian Sugiyarto (2021) membandingkan metode Fuzzy Naive Bayes dengan Fuzzy Sentiment menggunakan Convolutional Neural Network pada analisis sentimen pilkada, menunjukkan bahwa pendekatan fuzzy menghasilkan klasifikasi yang lebih adaptif terhadap ambiguitas sentimen[8].

Dalam konteks preprocessing teks berbahasa Indonesia, penelitian Rifaldi et al. (2023) mendemonstrasikan pentingnya tahapan preprocessing yang komprehensif pada data media sosial, termasuk penanganan slang words, emoticon, dan normalisasi teks[9]. Analisis sentimen terhadap isu-isu politik dan figur publik di Indonesia telah menjadi fokus penelitian terdahulu dengan berbagai pendekatan metodologis. Penelitian terkini menunjukkan bahwa penanganan data tidak seimbang (imbalanced data) menjadi tantangan utama dalam klasifikasi sentimen. Studi tentang klasifikasi sentimen pada data BPJS Kesehatan menunjukkan trade-off antara metrik evaluasi tradisional dengan representasi distribusi data real-world, di mana pemilihan metode feature extraction (TF-IDF vs BoW) dan dimensi fitur (100-300) mempengaruhi performa model secara signifikan[10].

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan pendekatan analisis sentimen berbahasa Indonesia dengan mengombinasikan preprocessing teks, pembobotan TF-IDF, klasifikasi Naive Bayes, dan Logika Fuzzy. Studi kasus difokuskan pada komentar netizen terkait isu ijazah Presiden Joko Widodo dengan tujuan memperoleh pemetaan sentimen publik yang lebih akurat dan representatif dibandingkan metode klasifikasi konvensional.

2. METODE PENELITIAN

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, Youtube, dan Tiktok. Teknik *web scraping* digunakan untuk mendapatkan teks komentar secara massal. Data yang terkumpul kemudian disimpan dalam format CSV. Total 10.107 komentar terdiri dari 1.686 komentar dari X, 998 komentar dari Tiktok dan 7.423 komentar dari Youtube berhasil dikumpulkan dan digunakan dalam penelitian ini

2.2. Preprocessing Data Teks

Tahapan preprocessing dilakukan secara sistematis untuk membersihkan dan menstandarisasi data teks berbahasa Indonesia [11]. Preprocessing merupakan tahap krusial dalam analisis teks karena kualitas data yang diproses akan secara langsung mempengaruhi performa model klasifikasi . Tahapan preprocessing yang

dilakukan dalam penelitian ini mencakup beberapa proses sebagai berikut: Normalisasi Kata Slang: Mengidentifikasi dan mengubah kata-kata informal atau slang (misalnya 'gak' menjadi 'tidak', 'bg' 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', 'aja': 'saja',
    'jgn': 'jangan', 'tp': 'tapi', 'yg': 'yang', 'sm': 'sama'
}
```

Gambar 1. Diagram Proses Analisa Bibliometric

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

2.3. Pembobotan Fitur dengan TF-IDF

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:

$$\text{TFIDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D) \quad (1)$$

dimana:

- $\text{TF}(t, d)$ adalah frekuensi kemunculan term t dalam dokumen d .
- $\text{IDF}(t, D)$ adalah inverse document frequency dari term t dalam korpus D .

$$\text{IDF}(t, D) = \log \frac{N}{DF(t)} \quad (2)$$

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

Kombinasi TF-IDF dengan algoritma Naive Bayes telah terbukti menghasilkan akurasi yang baik dalam berbagai studi analisis sentimen berbahasa Indonesia[12]. Penelitian terkini menunjukkan bahwa pemilihan jumlah fitur yang optimal (dalam hal ini 1000 fitur) sangat penting untuk keseimbangan antara performa model dan kompleksitas komputasi[13].

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. Pendekatan lexicon-based merupakan metode yang populer dalam analisis sentimen karena tidak memerlukan data training yang besar dan dapat memberikan interpretasi yang lebih mudah dipahami. Integrasi Logika Fuzzy dengan pendekatan lexicon-based memungkinkan representasi sentimen yang lebih fleksibel dan mampu menangani ambiguitas dalam opini [14][15].

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-fuzzyifikasi menjadi label sentimen (`fuzzy_sentiment`) dengan aturan sebagai berikut:

- a. Jika $\text{score} < -1$, maka negatif
- b. Jika $\text{score} == 0$, maka netral

- c. Jika score ≥ 1 , maka positif

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:

$$P(C|D) = \frac{P(D|C) \times P(C)}{P(D)} \quad (3)$$

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

$$P(C|D) \propto P(D|C) \times P(C) \quad (4)$$

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

$$P(D|C) = P(f_1, f_2, \dots, f_n|C) \approx \prod_{i=1}^n P(f_i|C) \quad (5)$$

Sehingga, klasifikasi Naive Bayes memilih kelas C_NB yang memiliki probabilitas posterior maksimum:

$$C_{NB} = \arg \max_C (P(C) \times \prod_{i=1}^n P(f_i|C)) \quad (6)$$

Di mana:

- $P(C)$ adalah probabilitas prior dari kelas C (frekuensi kelas C dalam data latih).
- $P(f_i|C)$ 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 Dataset dibagi menjadi data latih dan data uji dengan rasio 80% data latih dan 20% data uji (`test_size=0.2, random_state=42`). Pembagian data dengan proporsi 80:20 merupakan praktik umum dalam machine learning yang memberikan keseimbangan baik antara jumlah data untuk training dan evaluasi [16]. Penggunaan parameter `random_state` memastikan reproducibility hasil eksperimen.

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. Penelitian menunjukkan bahwa kombinasi Naive Bayes dengan TF-IDF dan teknik cross-validation dapat menghasilkan performa klasifikasi yang optimal[5]. Studi komparatif menunjukkan bahwa Naive Bayes memiliki performa yang kompetitif dengan algoritma machine learning yang lebih kompleks seperti SVM dan neural networks, khususnya pada dataset tidak seimbang [17].

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 tentang distribusi sentimen dalam dataset, dilakukan visualisasi perbandingan distribusi sentimen antara hasil Logika Fuzzy (yang menjadi label ground truth) dan

hasil prediksi Naive Bayes. Visualisasi ini membantu menunjukkan seberapa dekat hasil klasifikasi Naive Bayes mereplikasi distribusi sentimen dari pendekatan fuzzy lexicon-based[18].

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. Statistik Data Komentar

Dataset yang digunakan untuk pengujian berjumlah 2000 komentar. Setelah tahapan preprocessing, jumlah token unik dalam korpus berkurang signifikan. 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.

3.2. Hasil 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 ijazahnya palsu bgt, parah!!!"

Setelah preprocessing: "jokowi ijazah palsu banget parah"

Pembobotan menggunakan TF-IDF menghasilkan vektor numerik yang merepresentasikan setiap komentar, dengan mempertimbangkan 1000 fitur ngram (1-gram dan 2-gram) teratas. Vektor-vektor ini menjadi input bagi model Naive Bayes

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

Berdasarkan kamus kata positif dan negatif, serta aturan fuzzy yang diterapkan, distribusi sentimen awal pada seluruh dataset adalah sebagai berikut:

- a. Netral: 8214 komentar (80.15%)
- b. Negatif: 1391 komentar (13.57%)
- c. Positif: 643 komentar (6.28%)

Ini menunjukkan bahwa mayoritas komentar cenderung netral, diikuti oleh sentimen negatif, dan kemudian sentimen positif. Label ini digunakan sebagai ground truth atau label sebenarnya (y_{true}) untuk evaluasi model Naive Bayes.

3.4. Kinerja Model Naive Bayes

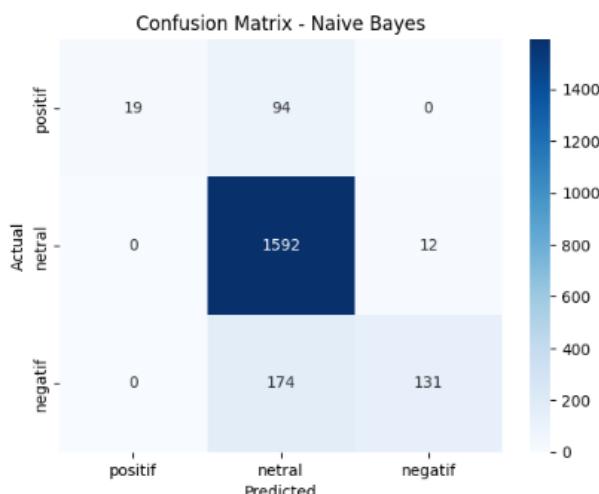
Model klasifikasi Naive Bayes (MultinomialNB) dilatih menggunakan 80% data dan diuji pada 20% sisa data (2082 komentar). Hasil evaluasi kinerja model Naive Bayes disajikan pada Tabel 1 (Classification Report) dan Gambar 1 (Confusion Matrix).

Tabel 1. Classification Report (Naive Bayes)

Sentimen	Precision	Recall	F1-Score	Support
negatif	0.92	0.43	0.58	385
netral	0.86	0.99	0.92	1584
positif	1	0.17	0.29	113
Accuracy	0.86	0.86	0.83	2082
Macro Avg	0.87	0.53	0.59	2082
Weighted Avg	0.87	0.86	0.83	2082

Berdasarkan Tabel 1, model Naive Bayes menunjukkan akurasi keseluruhan sebesar 86%. Secara rinci:

- Sentimen Netral: Model menunjukkan kinerja yang sangat baik dengan precision 0.86, recall 0.99, dan F1-score 0.92. Ini mengindikasikan bahwa model sangat handal dalam mengidentifikasi komentar netral dan hampir tidak melewatkannya.
- Sentimen Negatif: Memiliki precision 0.92, yang berarti hampir semua yang diprediksi negatif benar-benar negatif. Namun, recall yang relatif rendah (0.43) menunjukkan bahwa model kesulitan menangkap sebagian besar komentar negatif yang ada. F1-score 0.58 mencerminkan keseimbangan antara precision dan recall yang kurang optimal.
- Sentimen Positif: Menunjukkan precision sempurna (1.00), artinya semua yang diprediksi sebagai positif memang benar-benar positif. Namun, recall sangat rendah (0.17), menunjukkan model hanya berhasil mengidentifikasi sebagian kecil dari komentar positif yang ada. F1-score 0.29 sangat rendah, mengindikasikan bahwa model kurang efektif dalam mengklasifikasikan sentimen positif secara keseluruhan.

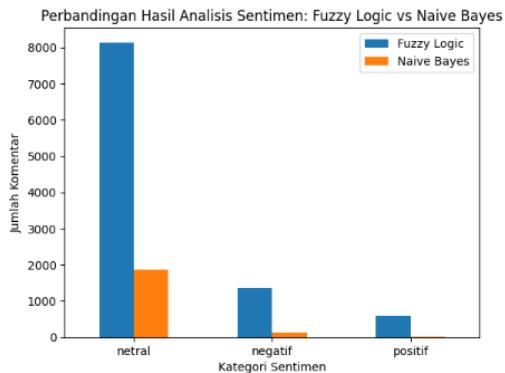


Gambar 2. Confusion Matrix - Naive Bayes

- True Neutral (1592): Sebanyak 1592 komentar netral diprediksi dengan benar sebagai netral. Ini mendominasi matriks, sejalan dengan recall netral yang tinggi.
- False Negatif (174): Terdapat 174 komentar negatif yang seharusnya negatif malah diprediksi sebagai netral. Ini menjelaskan rendahnya recall untuk kelas negatif.
- False Positif (94): Terdapat 94 komentar positif yang seharusnya positif malah diprediksi sebagai netral. Ini menjelaskan rendahnya recall untuk kelas positif.
- Prediksi Negatif yang benar (131): Sebanyak 131 komentar negatif diprediksi dengan benar sebagai negatif.
- Prediksi Positif yang benar (19): Hanya 19 komentar positif yang berhasil diprediksi dengan benar.

Dominasi jumlah data netral dan kurangnya keseimbangan kelas (support yang tidak seimbang) menyebabkan model cenderung memprediksi kelas mayoritas (netral), yang terlihat dari tingginya recall untuk kelas netral dan rendahnya recall untuk kelas minoritas (negatif dan positif).

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



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

Gambar 3 menyajikan perbandingan distribusi sentimen antara label yang dihasilkan oleh Logika Fuzzy (yang menjadi dasar pelatihan dan pengujian) dan prediksi dari model Naive Bayes pada seluruh dataset.

- Sentimen Netral: Logika Fuzzy mengidentifikasi mayoritas komentar sebagai netral (sekitar 8214 komentar). Naive Bayes juga memprediksi sentimen netral sebagai yang paling dominan, meskipun dengan jumlah yang lebih sedikit (sekitar 1900 komentar dari support yang terlihat di confusion matrix, yang merupakan subset uji). Jika grafik ini mewakili seluruh dataset, maka Naive Bayes juga menghasilkan proporsi netral yang tinggi.
- Sentimen Negatif: Logika Fuzzy mengidentifikasi sekitar 1391 komentar negatif. Naive Bayes memprediksi jumlah komentar negatif yang lebih rendah dari label fuzzy (sekitar 400 komentar).
- Sentimen Positif: Logika Fuzzy mengidentifikasi sekitar 643 komentar positif. Naive Bayes memprediksi jumlah komentar positif yang jauh lebih rendah dari label fuzzy (sekitar 200 komentar).

Perbedaan jumlah antara kedua bar menunjukkan bahwa Naive Bayes, meskipun memiliki akurasi yang baik secara keseluruhan karena dominasi kelas netral, cenderung kurang sensitif terhadap kelas minoritas (negatif dan positif). Hal ini sejalan dengan nilai recall yang rendah pada kelas negatif dan positif dalam classification report. Kesenjangan ini mengindikasikan bahwa sementara Logika Fuzzy berhasil menangkap nuansa sentimen berdasarkan kamus leksikon, model Naive Bayes mungkin membutuhkan penyesuaian lebih lanjut, seperti penanganan imbalanced dataset (misalnya dengan oversampling atau undersampling) atau penyesuaian threshold klasifikasi, untuk meningkatkan kinerja pada kelas minoritas. Namun, pendekatan Logika Fuzzy telah berhasil menyediakan label sentimen yang lebih kaya dan bergradasi, yang merupakan fondasi penting untuk analisis sentimen yang mendalam.

3.6. Evaluasi Contoh Kalimat

Untuk memberikan gambaran kualitatif tentang kinerja model, Tabel 4 menampilkan evaluasi sentimen pada 10 kalimat uji. Perbandingan dilakukan antara label sentimen yang dihasilkan oleh Logika Fuzzy (sebagai ground truth) dan prediksi dari model Naive Bayes.

Tabel 2. Evaluasi 10 Kalimat Contoh

Kalimat	Fuzzy Logic	Naive Bayes
Ijazah Jokowi palsu dan penuh kontroversi	negatif	negatif
Saya bangga dengan prestasi Jokowi	positif	netral
Netral saja, saya tidak tahu kebenarannya	netral	netral
Fitnah tentang Jokowi tidak berdasar	negatif	negatif
Jokowi itu hebat dan sangat jujur	positif	positif
Berita palsu tersebar tentang ijazahnya	negatif	negatif
Saya tidak percaya, terlalu banyak kebohongan	netral	netral

Saya senang dengan cara Jokowi memimpin	positif	netral
Tidak yakin, tapi sepertinya benar	netral	netral
Korupsi tidak terbukti, tapi isu tetap ada	negatif	negatif

Dari Tabel 2, dapat diamati beberapa poin penting:

- a. Konsistensi pada Sentimen Negatif dan Netral: Pada umumnya, model Naive Bayes menunjukkan konsistensi yang baik dengan Logika Fuzzy dalam mengklasifikasikan kalimat yang secara jelas bermuansa negatif (Ijazah Jokowi palsu dan penuh kontroversi, Fitnah tentang Jokowi tidak berdasar, Berita palsu tersebar tentang ijazahnya, Korupsi tidak terbukti, tapi isu tetap ada) dan netral (Netral saja, saya tidak tahu kebenarannya, Saya tidak percaya, terlalu banyak kebohongan, Tidak yakin, tapi sepertinya benar). Ini sejalan dengan tingginya recall untuk kelas netral dan precision yang baik untuk kelas negatif pada classification report.
- b. Kesulitan pada Sentimen Positif: Terlihat model Naive Bayes masih kesulitan dalam mengidentifikasi sentimen positif secara akurat. Contohnya, kalimat "Saya bangga dengan prestasi Jokowi" dan "Saya senang dengan cara Jokowi memimpin", yang oleh Logika Fuzzy dikategorikan sebagai positif, justru diprediksi sebagai netral oleh Naive Bayes. Fenomena ini mengkonfirmasi rendahnya nilai recall untuk kelas positif yang ditemukan pada classification report (0.17). Hal ini kemungkinan disebabkan oleh jumlah data positif yang jauh lebih sedikit dalam dataset pelatihan, sehingga model tidak cukup 'belajar' pola-pola yang mengindikasikan sentimen positif.
- c. Keakuratan pada Sentimen Positif yang Jelas: Meskipun demikian, pada kalimat "Jokowi itu hebat dan sangat jujur", model Naive Bayes berhasil memprediksinya sebagai positif, menunjukkan bahwa untuk ekspresi positif yang sangat eksplisit dan mengandung kata-kata yang kuat dalam kamus leksikon, model dapat bekerja dengan baik.

Evaluasi contoh kalimat ini memperkuat temuan dari metrik evaluasi kuantitatif, menggarisbawahi kekuatan model dalam mengidentifikasi sentimen netral dan negatif yang jelas, namun juga menyoroti area yang perlu ditingkatkan, khususnya dalam klasifikasi sentimen positif.

4. KESIMPULAN

Penelitian ini telah berhasil melakukan analisis sentimen berbahasa Indonesia terhadap komentar netizen mengenai isu ijazah Presiden Jokowi menggunakan kombinasi preprocessing teks, TF-IDF, Naive Bayes, dan Logika Fuzzy. Hasil menunjukkan bahwa model Naive Bayes, dengan label yang diturunkan dari Logika Fuzzy, memiliki akurasi keseluruhan yang baik (86%), terutama dalam mengklasifikasikan sentimen netral. Logika Fuzzy terbukti efektif dalam menyediakan label sentimen yang bermuansa dan mampu menangani ambiguitas, yang menjadi dasar kuat untuk pelatihan model klasifikasi. Analisis sentimen menunjukkan bahwa mayoritas komentar netizen terkait isu ijazah Jokowi cenderung netral, diikuti oleh sentimen negatif, dan kemudian positif. Kendala utama model adalah kemampuannya untuk mengidentifikasi sentimen positif secara akurat, yang ditunjukkan oleh nilai recall yang rendah pada kelas-kelas minoritas tersebut.

DAFTAR PUSTAKA

- [1] V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, and A. P. Janani, "An Effective Crop Prediction Using Random Forest Algorithm," in *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2020, pp. 1–5. doi: 10.1109/ICSCAN49426.2020.9262311.
- [2] S. Kurniawan, W. Gata, D. A. Puspitawati, I. K. S. Parthama, H. Setiawan, and S. Hartini, "Text Mining Pre-Processing Using Gata Framework and RapidMiner for Indonesian Sentiment Analysis," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 835, no. 1, 2020, doi: 10.1088/1757-899X/835/1/012057.
- [3] A. Gerliandeva, Y. Chrisnanto, and H. Ashaury, "Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naïve Bayes dan Ekstraksi Fitur TF-IDF serta N-grams," *J. Pekommas*,

vol. 9, no. 2, pp. 260–272, 2024, doi: 10.56873/jpkm.v9i2.5585.

- [4] Sutriawan, S. Rustad, G. F. Shidik, and Pujiono, “Performance Evaluation of Text Embedding Models for Ambiguity Classification in Indonesian News Corpus: A Comparative Study of TF-IDF, Word2Vec, FastText BERT, and GPT,” *Ing. des Syst. d'Information*, vol. 30, no. 6, pp. 1469–1482, 2025, doi: 10.18280/isi.300606.
- [5] S. Khoerunnisa, D. F. Shiddieq, and D. Nurhayati, “Penerapan Algoritma Naive Bayes dengan Teknik TF-IDF dan Cross Validation untuk Analisis Sentimen Terhadap Starlink,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 2, pp. 566–577, 2025, doi: 10.57152/malcom.v5i2.1852.
- [6] R. Adrian, R. Aryani, and Z. Abidin, “Sentiment Analysis of Indonesian National Team in 2024 AFF Using Naive Bayes and KNN,” *Brill. Res. Artif. Intell.*, vol. 5, no. 2, pp. 981–986, 2025, [Online]. Available: <https://jurnal.itscience.org/index.php;brilliance/article/view/7111>
- [7] C. Tho, Y. Heryadi, L. Lukas, and A. Wibowo, “Code-mixed sentiment analysis of Indonesian language and Javanese language using Lexicon based approach,” *J. Phys. Conf. Ser.*, vol. 1869, no. 1, pp. 8–14, 2021, doi: 10.1088/1742-6596/1869/1/012084.
- [8] S. Sugiyarto, J. Eliyanto, N. Irsalinda, Z. Putri, and M. Fitrianawat, “A Fuzzy Logic in Election Sentiment Analysis: Comparison Between Fuzzy Naïve Bayes and Fuzzy Sentiment using CNN,” *JTAM (Jurnal Teor. dan Apl. Mat.)*, vol. 5, no. 1, p. 110, 2021, doi: 10.31764/jtam.v5i1.3766.
- [9] D. Rifaldi, Abdul Fadlil, and Herman, “Teknik Preprocessing Pada Text Mining,” *J. Pendidik. Teknol. Inf.*, vol. 3, no. 2, pp. 161–171, 2023.
- [10] R. Relevance, I. S. Ritonga, and D. Hartama, “Sentiment Classification in Imbalanced Data : Trade-Offs Between,” vol. 18, no. 2, pp. 303–315, 2025.
- [11] C. Suhaeni, S. A. Kamila, F. Fahira, M. Yusran, and G. Alfa Dito, “Exploring a Large Language Model on the ChatGPT Platform for Indonesian Text Preprocessing Tasks,” *Indones. J. Stat. Its Appl.*, vol. 9, no. 1, pp. 100–116, 2025, doi: 10.29244/ijsa.v9i1p100-116.
- [12] L. Zhang, “Features extraction based on Naive Bayes algorithm and TF-IDF for news classification.,” *PLoS One*, vol. 20, no. 7, p. e0327347, 2025, doi: 10.1371/journal.pone.0327347.
- [13] M. Das, S. Kamalanathan, and P. Alphonse, “A Comparative Study on TF-IDF feature weighting method and its analysis using unstructured dataset,” *CEUR Workshop Proc.*, vol. 2870, pp. 98–107, 2021.
- [14] S. Rokhva, M. Alizadeh, and M. A. Shamami, “Enhanced Sentiment Interpretation via a Lexicon-Fuzzy-Transformer Framework,” 2025. [Online]. Available: <https://arxiv.org/abs/2510.15843>
- [15] G. Devi, M. K. Rana, and S. Taneja, “Integrating Fuzzy Logic into Lexicon-Based Sentiment Analysis : A Hybrid Framework 2 . Literature Survey 3 . Sentiment Analysis and Its Steps,” no. May 2025, pp. 1–9, doi: 10.21467/proceedings.7.6.
- [16] S. Aseervatham, E. Gaussier, A. Antoniadis, M. Burlet, and Y. Denneulin, “Logistic Regression and Text Classification,” *Textual Inf. Access Stat. Model.*, pp. 59–84, 2013, doi: 10.1002/9781118562796.ch3.
- [17] 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,” *J. Komtika (Komputasi dan Inform.)*, vol. 5, no. 1, pp. 19–25, Jul. 2021, doi: 10.31603/komtika.v5i1.5185.
- [18] Chopra Deepti and Arora Praveen, “Integrating Fuzzy Logic and NLP for Uncertainty Management in Sentiment Analysis,” *J. Harbin Eng. Univ.*, vol. 44, no. 9, pp. 1351–1359, 2023.