

Analisis Sentimen Program Makan Gratis Pada Platform X Menggunakan Algoritma Naïve Bayes

Methodius Modianus Laia¹, Fuad Nur Hasan², Antonius Yadi Kuntoro³

¹²³ Univeritas Bina Sarana Informatika, Jakarta Pusat, Jl. Kramat Raya No.98, Kwitang, Jakarta Pusat, 10450, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 21-07-2025

Revisi Akhir: 12-09-2025

Diterbitkan Online: 27-09-2025

KATA KUNCI

Sentiment analysis,

Naïve Bayes,

IndoBERT,

TF-IDF,

Platform X,

Free Meal Program.

KORESPONDENSI

E-mail: fuad.fnu@bsi.ac.id

ABSTRACT

The Free Meal Program is one of the government's strategic policies that has received various public responses, especially on social media Platform X (formerly Twitter). This study aims to analyze the level of public sentiment toward the Free Meal Program on Platform X. The classification method used is the Naïve Bayes algorithm, with model validation performed using the K-Fold Cross Validation technique. A total of 3,600 Indonesian-language tweets relevant to the Free Meal Program were collected through a web scraping process, followed by text preprocessing steps such as case folding, cleaning, tokenizing, stopword removal, and stemming. Data labeling was carried out semi-automatically using the IndoBERT model, and the tweets were then classified into two sentiment categories: positive and negative. The Naïve Bayes model was trained using the TF-IDF representation and tested on a test set comprising 20% of the total dataset. The evaluation results showed that the Naïve Bayes algorithm achieved an accuracy of 86.46%, precision of 86.55%, recall of 95.25%, and an F1-score of 90.77% on 458 test tweets. Validation using 10-fold cross-validation yielded an average accuracy of 86.74%. These results indicate that the Naïve Bayes algorithm demonstrates good classification performance and stable generalization in classifying public sentiment regarding the Free Meal Program. This research is expected to serve as a supporting tool in mapping public opinion based on social media.

1. PENDAHULUAN

Program Makan Gratis merupakan salah satu kebijakan unggulan yang diumumkan oleh calon presiden terpilih, Prabowo Subianto, dalam masa kampanye Pemilu 2024 [1]. Program ini telah mulai diimplementasikan secara bertahap sejak awal masa pemerintahan barunya pada Oktober 2024. Tujuan utama dari kebijakan ini adalah untuk menyediakan makanan bergizi secara gratis bagi siswa tingkat sekolah dasar hingga menengah, termasuk peserta didik di lingkungan pesantren. "Secara umum, program ini ditujukan untuk mengatasi permasalahan gizi serta meningkatkan kualitas sumber daya manusia di Indonesia, dengan fokus pada anak-anak dan ibu hamil" (Kevin Andreas Halomoan Tambunan et al., 2025) [2]

Meskipun memiliki tujuan yang positif, sejak diumumkan, program ini memicu berbagai perdebatan di tengah masyarakat. Methodius Modianus Laia

Isu terkait besarnya anggaran, kesiapan logistik, hingga potensi penyimpangan dalam pelaksanaan menjadi topik utama dalam diskusi publik. Perdebatan tersebut tidak hanya berlangsung di ruang diskusi formal, tetapi juga mencuat di dunia digital, khususnya melalui interaksi warganet di media sosial. Salah satu platform yang paling aktif digunakan dalam menyuarakan opini publik adalah Platform X (sebelumnya dikenal sebagai *Twitter*). "Sebagai media sosial berbasis teks dengan kecepatan penyebaran informasi yang tinggi, Platform X memiliki karakteristik yang relevan dalam mengamati dinamika opini masyarakat" (Muhammad et al., 2024) [3].

Dalam konteks program makan gratis, masyarakat mempunyai 2 pandangan yang berbeda, ada yang mendukung karena dinilai berpihak pada rakyat kecil, ada juga yang mengkritik karena khawatir akan ketidakefisienan pelaksanaan. Pandangan tersebut dibagikan dalam bentuk unggahan, komentar, hingga *thread* yang

viral [4]. Untuk mengetahui kecenderungan opini masyarakat secara sistematis, diperlukan pendekatan analisis sentimen.

“Analisis sentimen adalah proses mengidentifikasi dan mengkategorikan opini yang dinyatakan dalam teks, terutama untuk menentukan apakah sikap penulis terhadap topik tertentu bersifat positif, negatif, atau netral” (Ardiansyah et al., 2024) [5]. Dalam konteks ini, analisis sentimen berfungsi untuk memetakan persepsi publik terhadap program makan gratis melalui data komentar di Platform X.

Seiring perkembangan teknologi, analisis sentimen kini dapat dilakukan secara otomatis menggunakan pendekatan *machine learning*. “Pendekatan ini memungkinkan komputer untuk belajar dari data dan membuat klasifikasi sentimen secara efisien” (Hasibuan & Heriyanto, 2022). Salah satu algoritma yang banyak digunakan dalam tugas klasifikasi teks adalah *Naïve Bayes*. “Karena algoritma *Naïve Bayes* terkenal akan kemampuannya yang cukup baik dalam menangani data teks yang besar dan tidak terstruktur” (Aryanti & Santoso, 2023) [6].

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data primer yang diperoleh dari ulasan dan penilaian yang diberikan oleh pengguna” (Aldiansyah Pratama et al., 2024). Dengan mempertimbangkan relevansi sosial dari isu kebijakan ini serta kemampuan teknis dari metode yang digunakan, penelitian ini mengangkat topik “Analisis Sentimen Program Makan Gratis pada Platform X Menggunakan Algoritma *Naïve Bayes*” sebagai fokus kajian”.

2. TINJAUAN PUSTAKA

2.1 Program Makan Gratis

Program makan gratis merupakan program yang memberi makan dan susu secara gratis, untuk membantu pertumbuhan gigi anak dan ibu hamil yang merupakan program usulan dari presiden terpilih Prabowo Subianto dan wakil presiden Gibran Rakabuming Raka dalam masa pemerintahan 2024-2029 [7].

2.2 Media Sosial X (Twitter)

“X (atau sebelumnya *Twitter*) merupakan media sosial *microblogging* untuk mengekspresikan opini maupun perasaan penggunanya terhadap isu-isu atau permasalahan yang sedang maupun sudah terjadi” (Suparto & Habibullah, 2021) [8].

2.3 Analisis Sentimen

Analisis sentimen merupakan bidang kajian komputasional yang fokus untuk memahami, mengidentifikasi, dan mengekstraksi opini, emosi, atau sikap seseorang terhadap suatu entitas, baik itu produk, layanan, organisasi, atau

topik tertentu, yang umumnya diekspresikan dalam bentuk teks [9].

2.4 Algoritma *Naïve Bayes*

Naïve Bayes adalah algoritma pembelajaran yang sederhana dan efektif, berdasarkan aturan *Bayes* serta asumsi kuat bahwa tiap atribut bersifat kondisional independen terhadap kelas [10].

Rumus:

$$P(C_i | X) = \frac{P(X | C_i) \cdot P(C_i)}{P(X)}$$

Dimana :

C_i = kelas ke- i (misalnya *positif* atau *negatif*)

X = Data (fitur) yang akan diklasifikasikan

$P(C_i|X)$ = probabilitas suatu data termasuk ke kelas C_i

$P(X|C_i)$ = probabilitas data X muncul pada kelas C_i

$P(C_i)$ = probabilitas awal kelas C_i

$P(X)$ = probabilitas data X .

2.5 *K-Fold Cross Validation*

“*K-fold cross validation* merupakan cara statistikal yang digunakan dalam evaluasi performan dari model yang akan dibangun” (Wijiyanto et al., 2024). Dataset akan dibagi menggunakan 10-fold dengan ukuran yang sama pada setiap fold-nya, sehingga mempunyai 10 subset data.[11]

2.6 *IndoBERT*

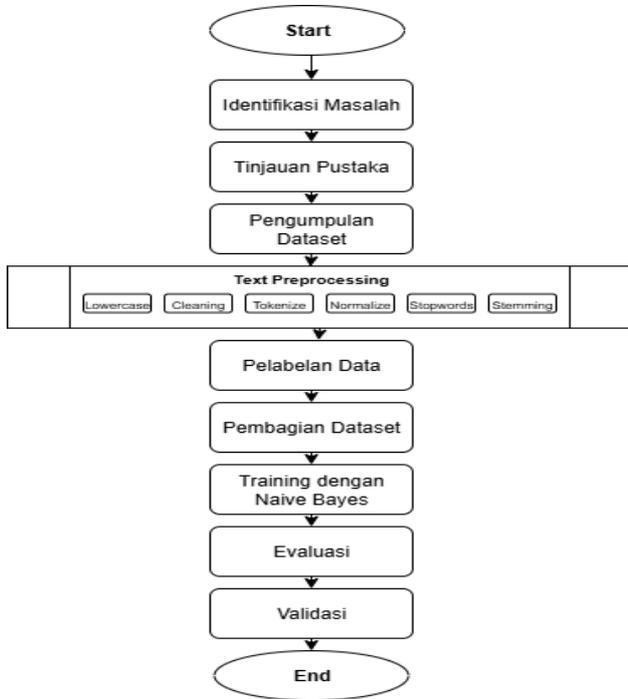
“*IndoBERT* merupakan model bahasa yang dikembangkan berdasarkan arsitektur *BERT* dan dilatih secara *monolingual* menggunakan korpus besar berbahasa Indonesia” (Koto et al., 2020) [12].

2.7 Penelitian Terdahulu

Penelitian yang paling relevan dengan studi ini adalah yang dilakukan oleh (Yoga Pratama et al., 2025) , yang secara eksplisit memanfaatkan model *IndoBERT*-base untuk klasifikasi sentimen terhadap Badan Pengelola Investasi Dan antara. Sebanyak 4.269 *tweet* dikumpulkan dan melalui proses pelabelan manual ke dalam tiga kategori: positif, negatif, dan netral. Dengan penerapan teknik augmentasi untuk mengatasi ketidakseimbangan data, model *IndoBERT* dilatih menggunakan pendekatan transfer *learning* dan berhasil mencapai akurasi sangat tinggi, yakni sebesar 97,71%. Temuan ini memperkuat efektivitas penggunaan model bahasa pra-latih berbasis transformer seperti *IndoBERT* dalam menangani tugas klasifikasi sentimen Bahasa Indonesia secara akurat dan efisien.

3 METODOLOGI

penelitian ini terdiri atas beberapa tahapan yang dilaksanakan secara sistematis, dimulai dari proses pengumpulan data hingga mencapai tahap akhir penyusunan hasil. Rangkaian tahapan tersebut ditampilkan secara visual dalam Gambar 1. sebagai alur penelitian.



Gambar. 1 Alur Penelitian

Data dikumpulkan menggunakan *tool* *TwitterHarvest*, sebuah alat *scraping* yang dapat mengakses data publik dari *X* secara efisien. Data dikumpulkan berdasarkan kata kunci spesifik yaitu: “MBG”, “Program makan gratis”, dan “Makan siang gratis”, dengan *filter* bahasa Indonesia (*lang:id*). Pengambilan dilakukan pada rentang waktu 1 Februari 2024 hingga 1 Juni 2025, yang mencakup masa kampanye, pelaksanaan, hingga reaksi publik pasca pemilu.

Metode penelitian yang digunakan dalam pengolahan dan analisis data adalah pendekatan klasifikasi teks berbasis *machine learning* dengan algoritma *Naïve Bayes*. Model ini digunakan untuk melakukan analisis sentimen terhadap data *tweet* yang telah dikumpulkan dan diproses sebelumnya. Setelah mempertimbangkan beberapa alternatif algoritma yang umum digunakan dalam klasifikasi teks, seperti *Support Vector Machine (SVM)* dan *Decision Tree*, algoritma *Naïve Bayes* dipilih karena memiliki keunggulan dari sisi efisiensi, kesederhanaan, dan performa yang cukup baik dalam menangani teks pendek seperti *tweet*.

Bahasa pemrograman *Python* digunakan untuk melakukan analisis data penelitian ini. Setelah pengujian model selesai, hasil analisis, yang terdiri dari jumlah pendapat yang dihasilkan dari analisis, akan divisualisasikan melalui aplikasi *Google Colabs*.

4 HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Data dikumpulkan dari platform *X* menggunakan *TwitterHarvest*. Proses *scraping* dilakukan dengan kata kunci: “MBG”, “Program makan gratis”, dan “Makan siang gratis” dengan *filter* bahasa Indonesia (*lang:id*) pada rentang waktu 1 Februari 2024 hingga 1 Juni 2025. *Scraping* dilakukan sebanyak 3 kali pada waktu berbeda mengingat penelitian ini menggunakan 3 kata kunci yang berbeda. *Dataset* disimpan dalam format *.csv* dan berisi kolom teks *tweet*, tanggal unggahan, *username*, dan meta data tambahan dapat dilihat pada Gambar dibawah ini.

```

# Crawl Data

filename = 'MBG1.csv'
search_keyword = 'MBG since:2024-02-01 until:2025-06-01 lang:id'
limit = 3600

!npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST" -l {limit} --token {twitter_auth_token}
    
```

Gambar. 2 Twitter Scapping dengan keyword MBG

```

# Crawl Data

filename = 'MBG1.csv'
search_keyword = 'Makan siang gratis since:2024-02-01 until:2025-06-01 lang:id'
limit = 1200

!npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST" -l {limit} --token {twitter_auth_token}
    
```

Gambar. 3 X scraping dengan keyword makan siang gratis

```

# Crawl Data

filename = 'MBG1.csv'
search_keyword = 'Program makan gratis since:2024-02-01 until:2025-06-01 lang:id'
limit = 1200

!npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST" -l {limit} --token {twitter_auth_token}
    
```

Gambar. 4 X scraping dengan keyword program makan gratis

	conversation_id_str	created_at	favorite_count	full_text	id_str	image_url
0	1928964524867985775	Sat May 31 23:39:15 +0000 2025	1	mbg ngabain anggaran doang anjrit pendidikan ...	1928964524867985775	NaN
1	1927667056482378065	Sat May 31 23:58:14 +0000 2025	0	@leannindonesia @EliasDipo @InvestorID @prabo...	1928964269216858378	NaN
2	1927667056482378065	Sat May 31 23:51:41 +0000 2025	0	@EliasDipo @InvestorID @prabowo Dari konsep ju...	1928962620607901880	NaN
3	1928022081196687377	Sat May 31 23:46:52 +0000 2025	0	@MasPe_22 Dukungan Presiden Prabowo jadi emerg...	1928961325293580309	NaN
4	1928021669701919214	Sat May 31 23:41:23 +0000 2025	0	@MasPe_22 Presiden Prabowo percaya pada potensi...	1928960027877949675	NaN

Gambar. 5 Hasil Scraping

4.2 Text preprocessing

Setelah data *tweet* terkumpul, tahap berikutnya yang sangat krusial adalah proses *text preprocessing*. Data yang diambil langsung dari *X* umumnya masih dalam bentuk mentah, mengandung banyak elemen yang tidak relevan atau bahkan mengganggu proses analisis, seperti URL, *mention* (@username), *hashtag*, emoji, angka, dan berbagai simbol. Oleh karena itu, dibutuhkan serangkaian tahap pembersihan informasi dalam teks menjadi lebih

```
df = dataset.loc[:, ['full_text']]
```

bermakna bagi model *machine learning* yang akan digunakan. Adapun untuk keperluan penelitian, dari dataset yang sudah dibentuk, hanya diambil kolom *full_text*, yaitu kumpulan *tweet*nya saja.

Gambar. 6 Pemisahan Kolom Dataset

Gambar. 7 Informasi tentang Dataset

```
<class 'pandas.core.frame.DataFrame'>
Index: 3600 entries, 0 to 3605
Data columns (total 1 columns):
# Column Non-Null Count Dtype
---
0 full_text 3600 non-null object
dtypes: object(1)
memory usage: 185.3+ KB
```

4.3 Pelabelan Data

Dalam penelitian ini, proses pelabelan dilakukan secara semi-otomatis menggunakan model *transformer* berbahasa Indonesia, yaitu *IndoBERT* yang telah dilatih khusus untuk tugas klasifikasi sentimen. Model ini dipilih karena memiliki kemampuan memahami konteks kalimat secara lebih mendalam, dibandingkan pendekatan rule-based atau leksikon biasa.

Setiap *tweet* yang telah dipreproses dimasukkan ke dalam pipeline prediksi model *IndoBERT*. Model akan mengeluarkan label sentimen yang diprediksi untuk masing-masing *tweet*. Hasil dari proses ini kemudian diverifikasi secara manual untuk memastikan akurasi

```
import transformers
import pandas as pd
from transformers import pipeline, AutoTokenizer
from tqdm.auto import tqdm

tqdm.pandas()

# Potong teks maksimal 300 kata
def truncate_text(text, max_words=300):
    return " ".join(text.split()[:max_words])

df['full_text'] = df['preprocessed'].astype(str).apply(truncate_text)

# Load classifier
classifier = pipeline(
    "sentiment-analysis",
    model="w1lw0/indonesian-roberta-base-sentiment-classifier",
    tokenizer="w1lw0/indonesian-roberta-base-sentiment-classifier",
    truncation=True,
    max_length=512
)

# Mapping label
def map_indobert(label):
    if label == 'positive':
        return 'positif'
    elif label == 'negative':
        return 'negatif'

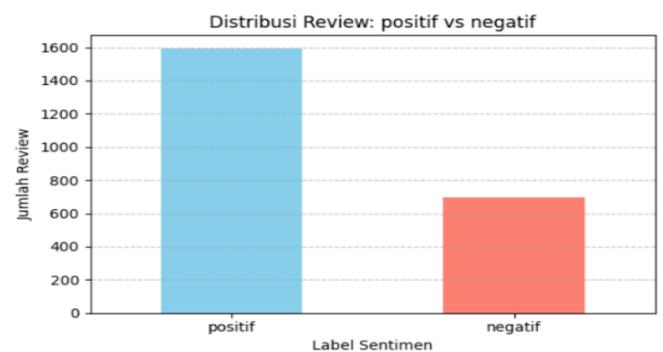
# Klasifikasi
df['label_indobert_raw'] = df['full_text'].progress_apply(lambda x: classifier(x))
df['label_indobert'] = df['label_indobert_raw'].apply(map_indobert)
```

pelabelan, terutama pada data yang memiliki ambiguitas atau nuansa sarkasme. Berikut merupakan cuplikan kode yang digunakan dalam tahap ini.

Gambar. 8 Penerapan Labelling

Tabel. 1 Hasil *labeling*

No.	<i>Tweet Asli</i>	<i>Tweet Preprocessed</i>	Label <i>Indobert</i>
1373	program makan gizi gratis mbg bawa ubah signif...	program makan gizi gratis mbg bawa ubah signif...	Positif
215	lah gapapaa kann nanti dapet mbg	lah gapapaa kann nanti dapet mbg	Negatif
844	tahun gizi anak bangsa jadi prioritas anggar p...	tahun gizi anak bangsa jadi prioritas anggar p...	Positif
671	mbg milik banyak manffat bagi masyarakat yaitu...	mbg milik banyak manffat bagi masyarakat yaitu...	Positif
2165	makan siang gratis alias wowo mau jual rakyat ...	makan siang gratis alias wowo mau jual rakyat ...	Negatif
1373	program makan gizi gratis mbg bawa ubah signif...	program makan gizi gratis mbg bawa ubah signif...	Positif



Gambar. 9 Distribusi Sentimen

4.4 Pembagian Dataset

Dataset yang telah diberi label kemudian dibagi ke dalam dua subset, yaitu:

1. Data latih (*training set*): digunakan untuk melatih model *Naïve Bayes*.
2. Data uji (*testing set*): digunakan untuk mengukur kemampuan generalisasi model terhadap data baru.

Pada penelitian ini, data dibagi dengan rasio 80% data latih dan 20% data uji. Berikut merupakan cuplikan kode yang digunakan dalam tahap ini.

```
# 3. Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42)
```

Gambar. 10 Splitting Dataset

Tabel. 2 Hasil *splitting* Dataset

Data	Jumlah <i>Tweet</i>	Persentase (%)
Data Latih	1829	80
Data Uji	458	20

4.5 Training Dengan *Naïve Bayes*

Sebelum proses pelatihan dimulai, seluruh teks yang telah melalui tahap *preprocessing* diubah terlebih dahulu menjadi bentuk representasi *numerik* agar dapat diproses oleh model pembelajaran mesin. Representasi yang digunakan dalam penelitian ini adalah *TF-IDF* (*Term Frequency-Inverse Document Frequency*).

Dengan *TF-IDF*, kata-kata yang sering muncul tetapi tidak terlalu umum dalam dataset akan diberi bobot yang lebih besar, sedangkan kata-kata umum yang muncul di hampir semua dokumen akan diredam bobotnya. Berikut merupakan cuplikan kode yang digunakan dalam tahap ini

```
# 2. TF-IDF dari teks yang sudah dibersihkan
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(df['cleaned_text']) # fitur teks
y = df['label_indobert'] # target label
```

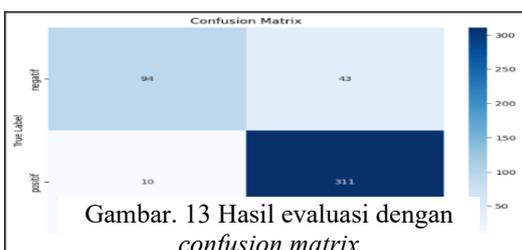
Gambar. 11 Kode Penerapan *TF-IDF*

```
# 4. Train model
model = MultinomialNB()
model.fit(X_train, y_train)
```

Gambar. 12 *Train Model*

4.6 Evaluasi

Evaluasi performa model *Naïve Bayes* dilakukan untuk mengukur seberapa baik model mampu mengklasifikasikan sentimen *tweet* ke dalam dua kategori, yaitu positif dan negatif. Proses evaluasi ini menggunakan data uji (*test set*) yang berjumlah 458 *tweet*, yaitu 20% dari total data yang telah diproses dan diberi label sebelumnya.



Gambar. 13 Hasil evaluasi dengan *confusion matrix*

4.7 Validasi

Model dievaluasi lebih lanjut menggunakan teknik *10-Fold Cross Validation* guna menguji stabilitas dan generalisasi model terhadap data yang berbeda.

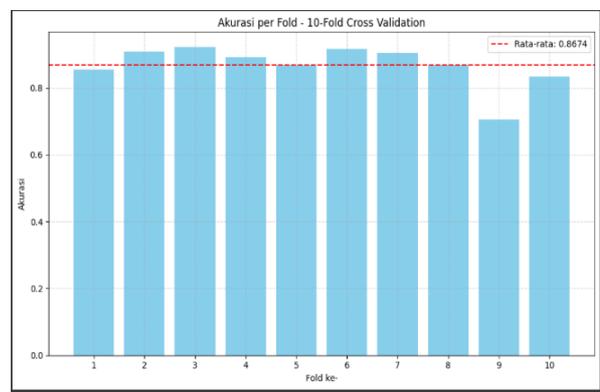
```
# 1. Load hasil preprocessing
df = pd.read_csv('Hasil_dengan_label.csv')
df.columns = df.columns.str.strip()
df['cleaned_text'] = df['preprocessed'].astype(str).fillna('')

# 2. TF-IDF
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(df['cleaned_text'])
y = df['label_indobert']

# 3. 10-fold cross-validation
model = MultinomialNB()
scores = cross_val_score(model, X, y, cv=10, scoring='accuracy')

print(f'Akurasi tiap fold: {scores}')
print(f'Rata-rata akurasi: {scores.mean():.4f}')
```

Gambar. 14 Penerapan *10-Fold Cross Validation*



Gambar. 15 Hasil evaluasi dengan *k-fold cross validation*

5 KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan mengenai analisis sentimen pengguna *X* terhadap program makan gratis menggunakan algoritma *Naïve Bayes*, diperoleh beberapa kesimpulan sebagai berikut:

1. Pengumpulan data *tweet* telah dilakukan melalui proses *web scraping* dengan kata kunci “MBG”, “Program makan gratis”, dan “Makan siang gratis” pada rentang waktu 1 Februari 2024 hingga 1 Juni 2025. Data kemudian diproses melalui tahapan standar NLP sebelum diberi label sentimen secara otomatis menggunakan *IndoBERT*.
2. Setelah dilakukan pelabelan dataset sebanyak 3600 *tweet* dengan menggunakan *IndoBERT* diketahui sebanyak 1591 *tweet* bernilai positif, sedangkan 696 *tweet* lainnya bernilai negatif, dan 1313 *tweet* bernilai netral. Karena kepentingan penelitian, *tweet* yang berlabel netral tidak digunakan.
3. Model klasifikasi *Naïve Bayes* dilatih menggunakan fitur representasi *TF-IDF* dan menunjukkan hasil yang baik dengan nilai akurasi sebesar 86,46%, *presisi* 86,55%, *recall* 95,25%, dan

F1-Score 90,77% terhadap data uji sebanyak 458 *tweet*.

4. Untuk menguji kestabilan performa model, dilakukan evaluasi tambahan menggunakan *10-Fold Cross Validation*. Hasilnya menunjukkan bahwa model memiliki rata-rata akurasi sebesar 86,74%, dengan sebagian besar *fold* menunjukkan akurasi di atas 86%. Hal ini mengindikasikan bahwa model memiliki kinerja yang konsisten dan dapat digeneralisasikan dengan baik pada data yang berbeda-beda, sehingga dapat diandalkan dalam tugas klasifikasi sentimen serupa di masa mendatang.
5. Berdasarkan hasil evaluasi model, maka hipotesis nol (H_0) yang menyatakan bahwa algoritma *Naïve Bayes* tidak mampu mengklasifikasikan sentimen masyarakat terhadap Program Makan Gratis secara akurat, ditolak.
6. Sebaliknya, hipotesis alternatif (H_1) yang menyatakan bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan sentimen masyarakat secara akurat, diterima, karena model terbukti menghasilkan performa klasifikasi yang tinggi dan dapat membedakan opini positif maupun negatif secara efektif.

DAFTAR PUSTAKA

- [1] Aldiansyah Pratama, M., Hasan, F. N., & Pratama, M. A. (2024). *Comparison of the Naïve Bayes Method and Support Vector Machine in Sentiment Analysis of Genshin Impact Game Reviews*. In *MECOMARE* (Vol. 13, Issue 2). www.trigin.pelnu.ac.id
- [2] Amelia Br Siregar, R., Ramadhan Manik, A., Syahri, A., Budi Akbar, M., Ramadhani, F., & Iskandar Psr V Medan Esatate Kab Deli Serdang, J.W. (2025). Penerapan *Naïve Bayes* Untuk Analisis Sentimen Netizen Terhadap Program Makan Siang Gratis Pada Media Sosial X. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Issue 4).
- [3] Ardelia Maharani, P., Riyani Namira, A., & Viony Chairunnisa, T. (2024). *Jolasos : Journal of Law and Social Society Peran Makan Siang Gratis Dalam Janji Kampanye Prabowo Gibran Dan Realisasinya*.
- [4] Ardiansyah, A., Argarini Pratama, E., Imam Fadlilah, N., & Bina Sarana Informatika, U. (2024). *Analisis Sentimen Pengguna Terhadap Aplikasi ChatGPT Di Google Play Store: Penerapan Algoritma Support Vector Machine*. *11(2)*, 247–254.
- [5] Aryanti, P. G., & Santoso, I. (2023). *Analisis Sentimen Pada Twitter Terhadap Mobil listrik Menggunakan Algoritma Naive Bayes*. <https://journals.upi.ac.id/index.php/ikraith-informatika/issue/archive>
- [6] Bing Liu. (2015). *Sentiment Analysis Mining Opinions, Sentiments, and Emotions*.
- [7] Gunawan, T. S., Ashraf, A., Riza, B. S., Haryanto, E. V., Rosnelly, R., Kartiwi, M., & Janin, Z. (2020). Development of video-based emotion recognition using deep learning with Google Colab. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18(5)*, 2463–2471. <https://doi.org/10.12928/TELKOMNIKA.v18i5.16717>
- [8] Hasibuan, E., & Heriyanto, E. A. (2022). Analisis sentimen pada ulasan aplikasi *amazon shopping di google play store* menggunakan *naive bayes classifier*. *Jts*, *1(3)*.
- [9] Kardandi Alfarizi Siregar. (2025). Analisis Sentimen Netizen Indonesia Terhadap Kampanye Penggunaan Kecerdasan Buatan Oleh Pemerintah Menggunakan Algoritma *Naive Bayes*.
- [10] Kevin Andreas Halomoan Tambunan, Ridha Nababan, Rimma Anisa Siagian, Roslin Naiborhu, Sintia Harianti, & Jamaludin Jamaludin. (2025). Tinjauan Kritis Tentang Program Makan Bergizi Gratis Terhadap Produktivitas Belajar Siswa. *Katalis Pendidikan : Jurnal Ilmu Pendidikan Dan Matematika*, *2(2)*, 21–31. <https://doi.org/10.62383/katalis.v2i2.1428>
- [11] Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). *IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP*. Online. <https://huggingface.co/>
- [12] Mitchell, Ryan. (2024). *Web scraping with Python : Data extraction from the modern web*. O'reilly Media.
- [13] Putriyeki, A., Sulistiawati, D., Khairani, N., Rizky, R., Kusuma, A., Keuangan, P., & Stan, N. (2025). Analisis Multidimensional Sentimen Masyarakat Terhadap Program Makan Bergizi Gratis Pada Media Sosial X. *Integrative Perspectives of Social and Science Journal*, *2(1)*, 1004.
- [14] Reynaldi Valerian, F., Syarief, M., Abdul Fatah, D., Raya Telang, J., Kamal, K., & Timur, J. (2025). Klasifikasi tingkat obesitas menggunakan metode *gbm* dan *confusion matrix*. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Issue 2).
- [15] Sakr, S., & Zomaya, A. Y. (2019). *Encyclopedia of Big Data Technologies*.
- [16] Sammut, C., & Webb, G. I. (2017). *Encyclopedia of Machine Learning and Data Mining Second Edition*.
- [17] Suparto, D., & Habibullah, A. (2021). Efektivitas Penggunaan Sosial Media *Twitter* dalam Penyebaran Informasi dalam Pelayanan Publik. *Indonesian Governance Journal : Kajian Politik-Pemerintahan*, *4(2)*, 161–172. <https://doi.org/10.24905/igi.v4i2.1927>

- [18] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma *Naïve Bayes* dan KNN. *Jurnal KomtekInfo*,1-7.
<https://doi.org/10.35134/komtekinfo.v10i1.330>
- [19] Wijiyanto, W., Pradana, A. I., Sopingi, S., & Atina, V. (2024). Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa. *Jurnal Algoritma*,21(1).
<https://doi.org/10.33364/algoritma/v.21-1.1618>
- [20] Yoga Pratama, A., Ananda Sanjaya, G., Khairunisa Lubis, N., & Rangga Aditya, M. (2025). *Analisis Sentimen Publik Terkait Danantara Menggunakan Algoritma IndoBERT pada Platform Media Sosial*. 9, 2025.
<https://doi.org/10.47002/metik.v9i1.1055>
- [21] Zheng, A. (2015). *Evaluating Machine Learning Models*.

BIODATA PENULIS



Methodius Modianus Laia

Merupakan Mahasiswa aktif di Universitas Bina Sarana Informatika



Fuad Nur Hasan, M.Kom

Merupakan Dosen aktif di Universitas Bina Sarana Informatika



Antonius Yadi Kuntoro, M.M, M.Kom

Merupakan Dosen aktif di Universitas Bina Sarana Informatika