



Analisis Sentimen Opini Publik Menggunakan Algoritma *Naive Bayes* dan TF-IDF

Yoga Handoko Agustin¹, Neng Cici Mulyani², Wahyu Sindu Prasetya³

^{1,2}Institut Teknologi Garut, Indonesia

³STMIK Pontianak, Indonesia

*email: yoga.handoko@itg.ac.id

Info Artikel

Dikirim: 12 Agustus 2025

Diterima: 25 Agustus 2025

Diterbitkan: 30 November 2025

Kata kunci:

Analisis Sentimen;

Instagram;

Kebijakan Study Tour;

Naive Bayes;

SEMMA.

ABSTRAK

Penelitian ini berfokus pada analisis sentimen masyarakat terhadap kebijakan larangan study tour yang dikeluarkan oleh Gubernur Jawa Barat dengan memanfaatkan data komentar dari media sosial Instagram. Data dikumpulkan melalui teknik web scraping menggunakan ekstensi Instant Data Scraper dengan kata kunci relevan, kemudian diberi label secara otomatis oleh ChatGPT. Untuk menjamin kualitas pelabelan, dilakukan validasi manual terhadap 10% data secara acak, yang menghasilkan tingkat akurasi sebesar 93%. Proses analisis dilakukan menggunakan algoritma *Naive Bayes* dengan kerangka kerja SEMMA (*Sample, Explore, Modify, Model, Assess*). Tantangan distribusi kelas yang tidak seimbang diatasi melalui penerapan SMOTE (*Synthetic Minority Over-sampling Technique*). Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan akurasi model sebesar 80%, dengan F1-score tertinggi pada kategori sentimen positif (82%) dan negatif (81%). Temuan ini membuktikan bahwa kombinasi SEMMA dan algoritma *Naive Bayes* efektif untuk memetakan opini publik berbasis data media sosial. Lebih jauh, penelitian ini memberikan kontribusi praktis bagi pemerintah dan pembuat kebijakan, khususnya dalam memonitor persepsi masyarakat secara real-time terhadap kebijakan yang diterapkan. Dengan pendekatan ini, pemerintah dapat lebih cepat mengidentifikasi respons publik, mengantisipasi potensi penolakan, serta menyusun strategi komunikasi yang lebih tepat sasaran. Selain itu, kerangka kerja yang digunakan dapat diadaptasi pada isu kebijakan lainnya, sehingga bermanfaat sebagai model analisis sentimen yang sistematis, terukur, dan mendukung pengambilan keputusan berbasis data.

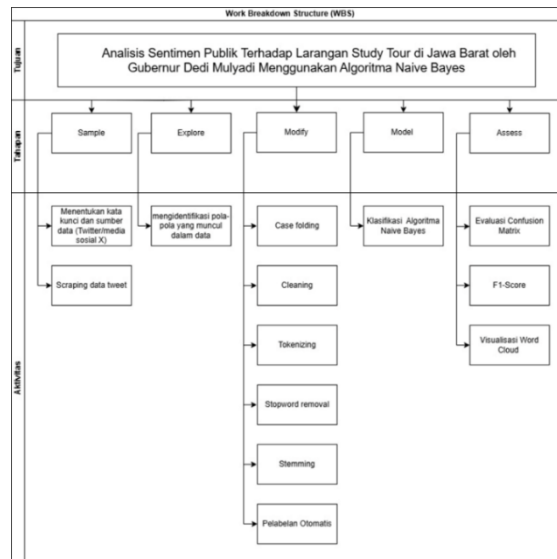
1. PENDAHULUAN

Media sosial telah berperan signifikan bagi masyarakat dalam mengekspresikan opini terhadap isu publik, termasuk kebijakan pemerintah [1]. Salah satu kebijakan yang menimbulkan reaksi beragam adalah larangan kegiatan study tour oleh Gubernur Jawa Barat. Study tour adalah kegiatan pembelajaran di luar kelas yang bertujuan memperkaya wawasan siswa melalui pengalaman langsung, sehingga materi dapat dikaitkan dengan realitas sosial, budaya, dan lingkungan kerja [2]. Melalui Instagram, publik menyampaikan pro dan kontra yang dapat dianalisis dengan metode analisis sentimen untuk memahami opini dan emosi masyarakat,

sehingga bermanfaat bagi pembuat kebijakan. [3]. Dalam konteks ini, algoritma *Naive Bayes* dipilih karena kemampuannya dalam menangani data teks yang besar dengan efisiensi komputasi yang tinggi [4]. Metodologi SEMMA (*Sample, Explore, Modify, Model, Assess*) digunakan untuk mengarahkan alur analisis secara sistematis [5]. Dengan pendekatan ini, penelitian diharapkan mampu menggambarkan persepsi masyarakat terhadap kebijakan pemerintah berbasis data sosial media. Analisis sentimen digunakan untuk mengidentifikasi dan mengelompokkan opini publik berdasarkan polaritasnya, seperti positif, negatif, dan netral sebagai metode yang efektif [6]. Dengan memanfaatkan algoritma pembelajaran mesin, khususnya *Naive Bayes*, proses klasifikasi sentimen dapat dilakukan secara sistematis dan efisien [7]. *Naive Bayes* dipilih karena kemampuannya dalam mengolah data teks berdimensi tinggi dengan asumsi independensi antar fitur, serta performanya yang baik pada dataset berskala besar [8]. Untuk menjaga keteraturan proses analisis, penelitian ini menerapkan kerangka kerja SEMMA (*Sample, Explore, Modify, Model, Assess*) yang dirancang untuk membantu peneliti melakukan pengolahan data secara bertahap [9]. Tahapan ini dimulai dari pengambilan sampel data, eksplorasi distribusi kelas, modifikasi melalui pembersihan dan penyeimbangan data, pembangunan model klasifikasi, hingga evaluasi hasil. Selain itu, untuk memastikan kualitas data latih, dilakukan pelabelan otomatis menggunakan model ChatGPT, yang kemudian divalidasi melalui cross-check manual pada 10% data yang diambil secara acak [10]. Distribusi kelas yang awalnya tidak seimbang diperbaiki menggunakan SMOTE (*Synthetic Minority Over-sampling Technique*) sehingga setiap kelas memiliki jumlah data yang setara.

2. METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja SEMMA (*Sample, Explore, Modify, Model, Assess*) yang dikembangkan oleh SAS Institute untuk memandu proses analisis data. Tahapan penelitian meliputi pengumpulan data, eksplorasi data, preprocessing data, pemodelan dan evaluasi hasil, serta visualisasi data dengan menggunakan word cloud. Tahapan SEMMA dalam penelitian ini disusun secara terstruktur dalam Work Breakdown Structure (WBS), yang ditampilkan pada Gambar 1.



Gambar 1. *Work Breakdown Structure*

Struktur kerja tersebut disusun berdasarkan lima tahapan utama dalam pendekatan SEMMA, yakni *Sample, Explore, Modify, Model, dan Assess*.

2.1 Sample

Tahap sample merupakan langkah awal pengambilan dan pemilihan data yang relevan dengan tujuan penelitian pengumpulan data komentar dari Instagram menggunakan metode crawling dengan kata kunci

tertentu yang berhubungan dengan kebijakan larangan study tour. Data dikumpulkan dalam format CSV untuk memudahkan pengolahan pada tahap berikutnya.

2.2 Explore

Tahap eksplorasi bertujuan untuk memahami karakteristik dataset. Analisis dilakukan terhadap jumlah komentar per kelas sentimen, panjang teks, serta distribusi kata. Hasil eksplorasi awal menunjukkan distribusi sentimen yang tidak seimbang, dengan jumlah komentar positif lebih banyak dibanding negatif dan netral. Visualisasi word cloud juga dibuat untuk melihat frekuensi kata-kata yang sering muncul pada masing-masing kategori sentimen.

2.3 Modify

Dalam penelitian ini, data opini publik dari Instagram diproses melalui serangkaian tahapan text preprocessing untuk memastikan kualitas data sebelum dilakukan analisis sentimen dengan algoritma *Naive Bayes* [11]. Proses preprocessing meliputi:

1) Case Folding

Pada tahap ini, semua huruf dalam teks diubah menjadi huruf kecil (*lowercase*) untuk menghilangkan perbedaan bentuk huruf yang tidak berpengaruh pada makna kata. Misalnya, kata “Setuju” dan “setuju” akan diperlakukan sebagai entitas yang sama. Langkah ini penting untuk memudahkan proses pencocokan kata pada tahap selanjutnya.

2) Cleaning

Aktivitas ini mencakup penghapusan elemen-elemen yang tidak relevan dalam teks, seperti tanda baca, angka, URL, *mention* (@user), hashtag, dan karakter khusus lainnya. *Cleaning* bertujuan untuk mengurangi gangguan dalam proses klasifikasi yang bisa disebabkan oleh informasi yang tidak membawa makna penting.

3) Tokenizing

Proses memecah teks atau kalimat menjadi bagian-bagian yang lebih kecil. Misalnya, kalimat “Saya setuju kebijakan ini” akan dipisah menjadi token [saya, setuju, kebijakan, ini]. Pemecahan ini memudahkan pengolahan data oleh algoritma, terutama dalam proses ekstraksi fitur.

4) Stopword Removal

Pada tahap ini, kata-kata umum (*stopwords*) yang memiliki frekuensi tinggi namun kontribusi semantiknya terhadap sentimen sangat kecil dihapus dari teks. Contohnya adalah kata “yang”, “dan”, “di”, “ke”, “untuk”, dan sebagainya. Tujuannya adalah memfokuskan analisis hanya pada kata-kata yang mengandung makna utama.

5) Stemming

Proses stemming mengubah kata turunan ke bentuk dasar (*root word*). Misalnya, kata “mendukung”, “didukung”, dan “pendukung” akan dikembalikan menjadi kata dasar “dukung”. Hal ini membantu menyederhanakan representasi teks dan meningkatkan akurasi klasifikasi..

6) Pelabelan Otomatis

Tahapan pelabelan otomatis untuk mengklasifikasikan seluruh dataset ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Proses pelabelan otomatis ini menggunakan ChatGPT sebagai model *Large Language Model (LLM)* yang mampu memahami konteks kalimat dan menentukan kelas sentimen secara cepat [12]. Untuk memastikan keakuratan hasil pelabelan otomatis, dilakukan validasi melalui cross-check manual terhadap 10% data yang dipilih secara acak dari keseluruhan dataset. Validasi ini bertujuan membandingkan kesesuaian antara label yang dihasilkan oleh sistem dengan label yang diberikan oleh peneliti.

Sebelum dilakukan pelatihan model, data terlebih dahulu ditangani dengan pendekatan SMOTE (*Synthetic Minority Over-sampling Technique*). Teknik ini digunakan untuk menyeimbangkan distribusi label sentimen karena ketidakseimbangan jumlah data pada masing-masing kelas dapat menurunkan kinerja model. Dalam penelitian [13] SMOTE diterapkan sebelum tahap pelatihan model untuk memastikan bahwa opini positif, negatif, dan netral mendapatkan porsi representasi yang setara. Metode ini terbukti mampu

meningkatkan akurasi, recall, dan *F1-score* [14], SMOTE menjadi salah satu pendekatan yang direkomendasikan dalam pengolahan data tidak seimbang dalam berbagai studi klasifikasi berbasis teks.

2.4 Model

Pada tahap ini dibangun model klasifikasi sentimen menggunakan algoritma *Naïve Bayes*. Model ini dipilih berdasarkan kemampuannya dalam memproses data teks berdimensi tinggi dengan waktu komputasi yang relatif cepat [15]. Dataset dibagi ke dalam subset pelatihan dan pengujian dengan rasio 80:20. Proses pelabelan sentimen dilakukan secara otomatis menggunakan ChatGPT, dengan kategori sentimen positif, negatif, dan netral.

2.5 Assess

Tahap evaluasi dilakukan menggunakan confusion matrix untuk menilai distribusi prediksi terhadap data aktual dan *F1-score* untuk mengukur keseimbangan presisi serta recall, sedangkan word cloud digunakan untuk memvisualisasikan kosakata dominan dalam opini publik.

3. HASIL DAN PEMBAHASAN

3.1 Sampel

Dataset yang digunakan dalam penelitian ini berupa komentar publik dari media sosial Instagram terkait kebijakan larangan study tour oleh Gubernur Jawa Barat dengan rentan waktu bulan Februari-April 2025. Data diperoleh melalui proses web crawling menggunakan kata kunci yang relevan, menghasilkan total 2.303 selanjutnya data di bersihkan melalui preprocessing sehingga data yang didapat berjumlah 1.871 komentar.

3.2 Explore

Tahap eksplorasi dilakukan untuk memahami karakteristik data yang diperoleh. Analisis awal mencakup distribusi frekuensi kata, panjang komentar, dan jumlah kemunculan kata kunci yang dominan pada setiap kategori sentimen. Selain itu, identifikasi kata yang sering muncul membantu dalam memahami persepsi publik, misalnya kata-kata seperti setuju, dukung, dan lanjutan sering muncul pada sentimen positif, sedangkan kata bebani, tidak setuju, dan hentikan dominan pada sentimen negative.

3.3 Modify

1) Case Folding

Mengubah seluruh huruf dalam teks menjadi huruf kecil (*lowercase*) untuk menyamakan representasi kata, sehingga “ Study”, “STUDY”, dan “study” dianggap sebagai entitas yang sama.

Tabel 1. Hasil *Case Folding*

text	→	text_casefold
Kayaknya ada yg salah dengan larangan study tour ke luar wilayah/provinsi. Yg harusnya dilarang itu pungutannya yg ugal2an, kewajiban ikut, dan klausul "ikut ga ikut tetep bayar". Mau tour di dalem kota pun kalau ugal2an kayak gitu, mendingan dilarang aja..	→	kayaknya ada yg salah dengan larangan study tour ke luar wilayah/provinsi. yg harusnya dilarang itu pungutannya yg ugal2an, kewajiban ikut, dan klausul "ikut ga ikut tetep bayar". mau tour di dalem kota pun kalau ugal2an kayak gitu, mendingan dilarang aja..

Seluruh karakter diubah menjadi huruf kecil tanpa mengubah tanda baca maupun ejaan, sebagai langkah normalisasi awal sebelum tokenisasi dan tahap pembersihan lanjutan.

2) Cleaning

Cleaning adalah tahap pembersihan teks dengan menghapus tanda baca, simbol, maupun karakter yang tidak relevan, sehingga data menjadi lebih rapi dan siap diproses pada tahap analisis berikutnya.

Tabel 2. Hasil Setelah *Cleaning*

text_casefold	→	text_clean
kayaknya ada yg salah dengan larangan study tour ke luar wilayah/provinsi. yg harusnya dilarang itu pungutannya yg ugal2an, kewajiban ikut, dan klausul "ikut ga ikut tetep bayar". mau tour di dalam kota pun kalau ugal2an kayak gitu, mendingan dilarang aja..	→	kayaknya ada yg salah dengan larangan study tour ke luar wilayahprovinsi yg harusnya dilarang itu pungutannya yg ugal2an kewajiban ikut dan klausul ikut ga ikut tetep bayarmau tour di dalam kota pun kalau ugal2an kayak gitu mendingan dilarang aja

Tahap text cleaning menghapus tanda baca dan karakter khusus agar data lebih bersih dan siap untuk tokenisasi

3) *Tokenizing*

Tokenisasi memecah text menjadi potongan-potongan kecil yang disebut token, tujuannya adalah untuk mengubah text panjang menjadi satuan yang lebih mudah dianalisis

Tabel 3. Hasil Setelah *Tokenizing*

text_clean	→	tokens
kayaknya ada yg salah dengan larangan study tour ke luar wilayahprovinsi yg harusnya dilarang itu pungutannya yg ugal2an kewajiban ikut dan klausul ikut ga ikut tetep bayar mau tour di dalam kota pun kalau ugal2an kayak gitu mendingan dilarang aja	→	['kayaknya', 'ada', 'yg', 'salah', 'dengan', 'larangan', 'study', 'tour', 'ke', 'luar', 'wilayahprovinsi', 'yg', 'harusnya', 'dilarang', 'itu', 'pungutannya', 'yg', 'kewajiban', 'ikut', 'dan', 'klausul', 'ikut', 'ga', 'ikut', 'tetep', 'bayar', 'mau', 'tour', 'di', 'dalam', 'kota', 'pun', 'kalau', 'kayak', 'gitu', 'mendingan', 'dilarang', 'aja']

Tabel 3. Menampilkan tahap tokenisasi, yaitu memecah teks yang sudah dibersihkan menjadi daftar kata (token) berdasarkan spasi.

4) *Stopword Removal*

Penghapusan kata-kata yang tidak mengandung makna khusus terhadap analisis sentimen, seperti “yang, dan, di”, dilakukan menggunakan daftar stopwords Bahasa Indonesia.

Tabel 4. Hasil Setelah *Stopword Removal*

tokens	→	tokens_no_stopwords
['kayaknya', 'ada', 'yg', 'salah', 'dengan', 'larangan', 'study', 'tour', 'ke', 'luar', 'wilayahprovinsi', 'yg', 'harusnya', 'dilarang', 'itu', 'pungutannya', 'yg', 'kewajiban', 'ikut', 'dan', 'klausul', 'ikut', 'ga', 'ikut', 'tetep', 'bayar', 'mau', 'tour', 'di', 'dalam', 'kota', 'pun', 'kalau', 'kayak', 'gitu', 'mendingan', 'dilarang', 'aja']	→	['salah', 'larangan', 'study', 'tour', 'wilayahprovinsi yg', 'dilarang', 'pungutannya', 'ugal2an', 'kewajiban', 'klausul', 'tetep', 'bayarmau', 'tour', 'dalam', 'kota', 'ugal2an', 'kayak', 'gitu', 'mendingan', 'dilarang']

Tabel 4. Menunjukkan tahap stopwords removal, yaitu menghapus kata-kata umum yang tidak memiliki makna penting (misalnya: ada, yg, dengan, ke, itu, dan, di, pun, kalau). Hasilnya, hanya kata-kata bermakna seperti “salah, larangan, study, tour, pungutannya, kewajiban” yang dipertahankan, sehingga teks lebih fokus untuk analisis sentimen.

5) *Stemming*

Tahap stemming yaitu mengubah kata-kata ke bentuk dasar. Misalnya kata “dihentikan” diubah menjadi “henti”.

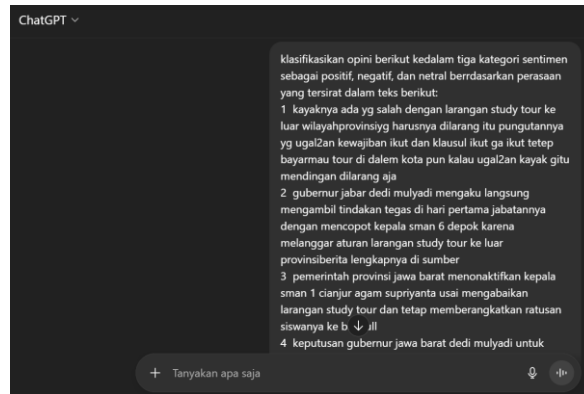
Tabel 5. Hasil Setelah *Stemming*

tokens_no_stopwords	→	tokens_stemmed
['salah', 'larangan', 'study', 'tour', 'wilayahprovinsi yg', 'dilarang', 'pungutannya', 'ugal2an', 'kewajiban', 'klausul', 'tetep', 'bayarmau', 'tour', 'dalam', 'kota', 'ugal2an', 'kayak', 'gitu', 'mendingan', 'dilarang']	→	['salah', 'larang', 'study', 'tour', 'wilayahprovinsi yg', 'larang', 'pungut', 'ugal2an', 'wajib', 'klausul', 'tetep', 'bayarmau', 'tour', 'dalam', 'kota', 'ugal2an', 'kayak', 'gitu', 'mending', 'larang']

Tabel 5. Menunjukkan tahap stemming, yaitu mengubah kata ke bentuk dasar. Misalnya “larangan” dan “dilarang” menjadi “larang”, “kewajiban” menjadi “wajib”, dan “mendingan” menjadi “mending”. Tahap ini bertujuan menyatukan variasi kata agar analisis sentimen lebih konsisten dan akurat.

6) Pelabelan otomatis

Pelabelan otomatis menggunakan ChatGPT, yang dimanfaatkan untuk mengklasifikasikan komentar ke dalam tiga kategori (positif, negatif, netral) berdasarkan konteks linguistik. ChatGPT dipilih karena kemampuannya memahami bahasa alami dan menghasilkan klasifikasi secara cepat serta konsisten. Untuk menjaga transparansi, model diberikan instruksi berbentuk prompt terstruktur. Berikut contoh prompt yang digunakan:

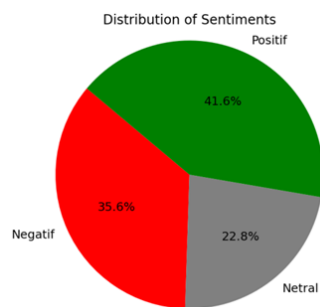


Gambar 2. Prompt Pelabelan Otomatis

Melalui pendekatan ini, setiap komentar diproses secara otomatis sehingga proses pelabelan menjadi lebih cepat dan efisien

Tabel 6. Hasil Pelabelan Dengan ChatGPT

No	text	label
1	kayaknya ada yg salah dengan larangan study tour ke luar wilayahprovinsi yg harusnya dilarang itu pungutannya yg ugal2an kewajiban ikut dan klausul ikut ga ikut tetep bayarmau tour di dalem kota pun kalau ugal2an kayak gitu mendingan dilarang aja	Negatif
2	gubernur jabar dedi mulyadi mengaku langsung mengambil tindakan tegas di hari pertama jabatannya dengan mencopot kepala sman 6 depok karena melanggar aturan larangan study tour ke luar provinsiberita lengkapnya di sumber	Netral
3	pemerintah provinsi jawa barat menonaktifkan kepala sman 1 cianjur agam supriyanta usai mengabaikan larangan study tour dan tetap memberangkatkan ratusan siswanya ke bali full	Netral
...
1871	kalo di sekolahku gak wajib sih	Netral



Gambar 3. Distribusi Sentimen

Distribusi sentimen menunjukkan bahwa opini publik cenderung terpolarisasi, dengan 41,6% komentar positif dan 35,6% komentar negatif, sedangkan komentar netral hanya 22,8%. Hal ini mengindikasikan bahwa kebijakan larangan study tour memicu respons emosional yang cukup kuat, baik berupa dukungan maupun penolakan, sementara respon netral relatif lebih sedikit. Untuk memastikan keakuratan pelabelan otomatis yang dilakukan oleh model ChatGPT dalam penelitian ini, dilakukan validasi melalui proses cross-check manual terhadap 10% data yang diambil secara acak dari keseluruhan dataset. Proporsi ini dipilih sebagai pertimbangan keseimbangan antara keterbatasan sumber daya dan kebutuhan memastikan kualitas pelabelan, serta mencerminkan metode validasi sampel yang banyak diterapkan pada penelitian berbasis machine learning.

Tabel 7. Perbandingan Label Otomatis dan Label Manual

No	text	label_otomatis	label_manual
1	saya setuju pak	Positif	Positif
2	kami ortu dukung kdm	Positif	Positif
3	smoga jateng jga ada yg speak up buat di hentikan study tour n wisuda n iuran kain2 yg membebani orang tua	Negatif	Positif
...
188	setiap kebijakan pasti ada aja pro kontra nya tetap lanjutkan pak	Positif	Positif

Tabel 8. Hasil Akurasi Pelabelan Otomatis

Label Manual	Label Otomatis			Precision	F1-Score
	Negatif	Netral	Positif		
Negatif	59	8	5	0.82	0.90
Netral	0	36	0	1.00	0.89
Positif	0	1	78	0.99	0.96
Recall	1.00	0.80	0.94		
Accuracy		93%			

Berdasarkan Tabel hasil perbandingan label, kategori sentimen negatif menunjukkan precision sebesar 0.82, recall 1.00, dan F1-score 0.90, menandakan bahwa semua opini yang tergolong negatif berhasil terdeteksi dengan sempurna, meskipun masih ada beberapa opini lain yang salah diklasifikasikan sebagai negatif. Pada kategori netral, precision-nya mencapai 1.00, dengan recall 0.80 dan F1-score 0.89, mengindikasikan bahwa meskipun semua prediksi netral tepat, masih terdapat opini netral yang gagal dikenali. Sementara itu, kategori positif menunjukkan performa terbaik dengan precision 0.99, recall 0.94, dan F1-score 0.96, yang berarti model dapat bekerja dengan akurat dalam mengklasifikasikan opini positif.

$$Accuracy = (59 + 36 + 78) / 187 + 173 / 187 = 0,9251 \approx 93\% \tag{1}$$

Hasil evaluasi menunjukkan bahwa pelabelan otomatis memiliki akurasi keseluruhan sebesar 93%, yang mencerminkan tingkat kesesuaian yang sangat tinggi dengan hasil pelabelan manual. Temuan ini menunjukkan bahwa pelabelan otomatis dengan ChatGPT dalam konteks analisis sentimen terhadap opini publik mengenai kebijakan larangan study tour dapat diandalkan, dan konsisten dengan penilaian manusia. Sebelum dilakukan pelatihan model, data terlebih dahulu ditangani dengan pendekatan SMOTE (*Synthetic Minority Over-sampling Technique*), karena kondisi distribusi data pada tahap awal menunjukkan ketidakseimbangan, di mana kelas positif lebih dominan dibandingkan negatif dan netral. Situasi ini berpotensi menyebabkan bias model terhadap kelas mayoritas, sehingga dilakukan penyeimbangan menggunakan SMOTE (*Synthetic Minority Over-sampling Technique*). Teknik ini terbukti meningkatkan kinerja model karena setiap kelas memperoleh representasi yang lebih adil.

Tabel 9. Distribusi Data Sebelum SMOTE

Label	Jumlah Data
Positif	778
Negatif	666
Netral	427

Data opini publik mengenai kebijakan larangan study tour didominasi oleh kelas positif sebanyak 778 data, diikuti oleh kelas negatif sebanyak 666 data, dan yang paling sedikit adalah sentimen netral sebanyak 427 data dengan total data 1.871. Untuk mengatasi permasalahan tersebut, dilakukan proses oversampling menggunakan metode SMOTE yang secara sintetik menghasilkan data baru pada kelas yang kekurangan. Setelah dilakukan proses ini, jumlah data di setiap kelas menjadi seimbang.

Tabel 10. Distribusi Data Setelah SMOTE

Label	Jumlah Data
Positif	778
Negatif	778
Netral	778

Penyeimbangan ini terbukti meningkatkan performa model, karena setiap kelas memperoleh representasi yang lebih proporsional dalam proses pembelajaran. Hasilnya, masing-masing kelas disamakan menjadi 778 data, sehingga total data menjadi 2.334 entri yang seimbang. Proses ini bertujuan agar model *Naive Bayes* tidak bias terhadap kelas mayoritas dan dapat melakukan klasifikasi secara lebih adil dan akurat.

3.4 Pemodelan Algoritma *Naive Bayes*

Tahap model bertujuan membangun sistem klasifikasi dengan memanfaatkan algoritma *Naive Bayes*. Model ini digunakan karena kemampuannya yang andal dalam menangani data teks pendek, perhitungan probabilitas yang efisien, dan kinerja yang baik pada data dengan jumlah fitur yang besar. Dataset hasil preprocessing dan SMOTE dibagi menjadi data latih dan data uji, kemudian digunakan untuk membangun model klasifikasi tiga kelas sentiment dengan menuliskan kode program di Google Colab yang memanfaatkan Python sebagai bahasa pemrograman.

```

1 import pandas as pd
2 from sklearn.model_selection import train_test_split
3 from sklearn.feature_extraction.text import TfidfVectorizer
4 from sklearn.naive_bayes import MultinomialNB
5 from sklearn.metrics import classification_report, accuracy_score, confusion_matrix
6 from imblearn.over_sampling import SMOTE
7 from collections import Counter
8
9 # LOAD DATA
10 df = pd.read_csv('/content/StudyTour_Label.csv')
11
12 # Split Data
13 X = df['text_clean']
14 y = df['label']
15
16 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
17
18 # TF-IDF vectorizer
19 tfidf = TfidfVectorizer()
20 X_tfidf = tfidf.fit_transform(X)
21
22 # SMOTE
23 smote = SMOTE(random_state=42)
24 X_resampled, y_resampled = smote.fit_resample(X_tfidf, y)
25
26 # Cek distribusi label setelah SMOTE
27 print("Distribusi label setelah SMOTE:", Counter(y_resampled))
28
29 # Split data
30 X_train, X_test, y_train, y_test = train_test_split(X_resampled, y_resampled, test_size=0.2, random_state=42)
31
32 # Train Naive Bayes
33 model = MultinomialNB()
34 model.fit(X_train, y_train)
35
36 # Prediksi dan Evaluasi
37 y_pred = model.predict(X_test)
38
39 # Evaluasi Model
40 print("Akurasi:", accuracy_score(y_test, y_pred))
41 print("\nClassification Report:\n", classification_report(y_test, y_pred))
42 print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))

```

Gambar 4. Source Code Pemodelan *Naive Bayes*

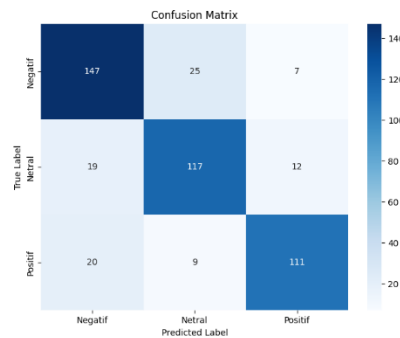
Gambar 4. menunjukkan proses klasifikasi sentimen otomatis menggunakan model *Naive Bayes* dalam bahasa Python dengan bantuan berbagai pustaka dari scikit-learn dan imbalanced-learn. Pertama-tama, data dimuat dari file `StudyTour_Label.csv`, kemudian dipisahkan menjadi dua bagian utama: fitur teks (`text_clean`) sebagai variabel input dan label sentimen (`label`) sebagai target. Setelah itu, data teks diubah ke bentuk numerik dengan bantuan teknik TF-IDF (Term Frequency-Inverse Document Frequency) agar bisa diproses oleh algoritma pembelajaran mesin. Karena data sentimen seringkali tidak seimbang (jumlah label positif, negatif, dan netral bisa sangat berbeda), kode ini menggunakan teknik SMOTE (Synthetic Minority Oversampling Technique) agar distribusi data menjadi seimbang di setiap kelas dengan membangun data

sintetis bagi kategori minoritas. Setelah penyeimbangan dilakukan, data kembali dibagi menjadi data latih dan data uji. Selanjutnya, model Multinomial *Naive Bayes* dilakukan pelatihan berdasarkan data latih, dan dilakukan prediksi terhadap data uji.

3.5 Assess

Setelah algoritma *Naive Bayes* diterapkan pada data komentar terkait kebijakan larangan study tour, tahap berikutnya yakni melakukan pengujian terhadap performa model klasifikasi yang telah dibuat. Evaluasi model dilakukan dengan memanfaatkan *confusion matrix* dengan indikator *accuracy*, *precision*, *recall*, *F1-score* dan visualisasi *word cloud*

1) Evaluasi Confusion Matrix



Gambar 5. Confusion Matrix

Gambar 5 merupakan confusion matrix yang digunakan untuk mengevaluasi kinerja model klasifikasi sentimen dalam tiga kategori, yaitu Negatif, Netral, dan Positif. Matriks ini mengilustrasikan bahwa model dapat mengklasifikasikan data secara akurat, berdasarkan hasil evaluasi model menggunakan confusion matrix, model berhasil mengklasifikasikan 147 data sebagai negatif, 117 data sebagai netral, dan 111 data sebagai positif secara benar. Jumlah prediksi benar tersebut berjumlah 375 dari total 467 data uji. Oleh karena itu, nilai akurasi model dihitung dengan rumus berikut:

$$Accuracy = \frac{147+117+111}{467} + \frac{375}{467} = 0,8029 \approx 80\% \tag{2}$$

Tabel 11. Evaluasi Kinerja Model

Aktual	Prediksi			Recall	F1-Score
	Negatif	Netral	Positif		
Negatif	147	25	7	82%	81%
Netral	19	117	12	79%	78%
Positif	20	9	111	79%	82%
Precision	79%	77%	85%		

Pada kelas negatif, model berhasil mengklasifikasikan 147 data dengan benar dengan precision 79%, recall 82%, dan F1-score 81%, yang menunjukkan keseimbangan antara ketepatan dan kemampuan deteksi. Pada kelas netral, performa relatif lebih rendah dengan precision 77%, recall 79%, dan F1-score 78%, mengindikasikan adanya kesulitan model dalam mengenali sentimen netral secara konsisten. Sementara itu, kelas positif menunjukkan hasil terbaik dengan precision 85%, recall 79%, dan F1-score 82%, yang menegaskan kekuatan model dalam mendeteksi opini positif secara akurat dan seimbang. Hasil penelitian ini menunjukkan bahwa algoritma *Naive Bayes* mampu mencapai akurasi sebesar 80%, Jika dibandingkan dengan penelitian [16] yang menganalisis opini masyarakat terhadap kebijakan vaksinasi booster menggunakan algoritma yang sama dengan akurasi 86,84%, maka capaian penelitian ini masih tergolong kompetitif meskipun sedikit lebih rendah. Perbedaan hasil dapat dipengaruhi oleh karakteristik isu kebijakan yang diteliti, variasi dataset, serta teknik preprocessing yang diterapkan. Hal

Cloud ini mencerminkan adanya kekhawatiran, ketidakpuasan, serta kritik dari masyarakat terhadap kebijakan larangan tersebut, khususnya karena dianggap menghambat kegiatan pembelajaran luar kelas yang bermakna.



Gambar 8. Word Cloud Sentimen Netral

Gambar 8 menampilkan visualisasi word cloud dari sentimen netral terhadap larangan study tour di Jawa Barat. Terdapat kata yang paling sering muncul antara lain "sekolah", "study", "tour", "anak", "jabar", dan "gubernur", yang menandakan fokus pembahasan publik berkisar pada topik utama kebijakan tersebut. Selain itu, kata "dilarang", "dedi", "mulyadi", "wisuda", dan "perpisahan" juga banyak digunakan, menggambarkan narasi informatif yang menyampaikan fakta atau opini yang tidak menunjukkan kecenderungan ke arah positif maupun negatif. Kehadiran kata-kata seperti "kang", "provinsi", "surat edaran", dan "kegiatan" mengindikasikan bahwa komentar dalam kategori ini cenderung mendeskripsikan konteks kebijakan atau sekadar menyampaikan kembali informasi yang telah beredar tanpa ekspresi emosi tertentu. Ini menunjukkan bahwa masyarakat juga menyampaikan informasi dan pendapat secara faktual, baik berupa pertanyaan, penjelasan, atau deskripsi situasi tanpa memuat ekspresi emosi yang kuat.

Hasil penelitian ini, dapat disimpulkan bahwa analisis sentimen publik terhadap larangan study tour di Jawa Barat memberikan gambaran yang jelas mengenai persepsi masyarakat, baik positif, negatif, maupun netral. Penelitian ini menunjukkan bahwa metode analisis sentimen dapat menjadi alat yang efektif untuk memantau opini publik secara real-time, membantu pemerintah dan lembaga terkait dalam merumuskan strategi komunikasi yang lebih tepat, serta mendukung pengambilan keputusan berbasis data. Kontribusi praktis dari penelitian ini terletak pada kemampuannya untuk menjadi acuan bagi pengambil kebijakan dan pemangku kepentingan lain, termasuk akademisi dan media sosial, dalam memahami tren opini masyarakat serta menyesuaikan kebijakan agar lebih responsif dan diterima oleh publik.

4. KESIMPULAN

Hasil pengujian menunjukkan algoritma *Naïve Bayes* mampu mengklasifikasikan komentar ke dalam tiga kategori sentimen dengan akurasi keseluruhan sebesar 80%. Nilai *F1-score* tertinggi diperoleh pada kategori positif (82%) dan negatif (81%), sedangkan kategori netral memperoleh *F1-score* sebesar 78%. Validasi pelabelan otomatis dengan cross-check manual terhadap 10% sampel acak dari dataset menghasilkan akurasi 93%, yang menunjukkan tingkat konsistensi tinggi pada metode pelabelan otomatis, meskipun validasi manual tetap disarankan untuk memastikan keakuratan. Temuan ini membuktikan bahwa pendekatan berbasis *Naïve Bayes* yang dipadukan dengan SMOTE dan metodologi SEMMA dapat menjadi strategi efektif untuk analisis sentimen berbasis media sosial, khususnya dalam konteks kebijakan publik, serta memiliki potensi besar untuk dimanfaatkan pengambil kebijakan dalam memantau opini publik secara real-time dan menyesuaikan strategi komunikasi maupun keputusan kebijakan.

Berdasarkan hasil penelitian, beberapa rekomendasi untuk penelitian selanjutnya antara lain: (1) Membandingkan performa dengan algoritma lain seperti Support Vector Machine (SVM), Random Forest, atau model deep learning seperti BERT; (2) Mengoptimalkan pelabelan otomatis dengan teknik prompt engineering dan model AI yang lebih terlatih khusus untuk Bahasa Indonesia; dan (3) Memperluas sumber

data dari media sosial lain seperti Twitter, Facebook, atau TikTok untuk memperoleh representasi opini publik yang lebih beragam. Dengan mengacu pada saran tersebut, penelitian di masa mendatang diharapkan dapat menghasilkan sistem analisis sentimen yang lebih akurat, adaptif, dan bermanfaat bagi pengambil kebijakan, lembaga pemerintah, maupun organisasi lain yang ingin memahami opini publik secara lebih efektif.

REFERENSI

- [1] A. Turmudi Zy, A. Nugroho, A. Rivaldi, and I. Afriantoro, “Analisis Sentimen Terhadap Pembobolan Data pada Twitter dengan Algoritma *Naive Bayes*,” *J. Teknol. Inform. dan Komput.*, vol. 8, no. 2, pp. 202–213, 2022, doi: 10.37012/jtik.v8i2.1240.
- [2] A. Rivaldy, H. A. Fedria Wowor, S. R. Maisya, and D. Safitri, “Penggunaan Twitter Dalam Meningkatkan Melek Politik Mahasiswa Ilmu Komunikasi Universitas Negeri Jakarta,” *Perspekt. Komun. J. Ilmu Komun. Polit. dan Komun. Bisnis*, vol. 5, no. 1, p. 41, 2021, doi: 10.24853/pk.5.1.41-48.
- [3] S. Andini, R. Kurniawan, S. Anwar, and K. Cirebon, “Analisis Sentimen Pengguna X Mengenai Opini,” vol. 13, no. 2, pp. 665–671, 2025.
- [4] M. I. Fikri, T. S. Sabrila, and Y. Azhar, “Comparison of *Naive Bayes* and Support Vector Machine Methods in Twitter Sentiment Analysis,” *Smatika J.*, vol. 10, no. 02, pp. 71–76, 2020.
- [5] SAS, “SAS ® Enterprise Miner™ 15.2: Reference Help SAS ® Documentation,” p. 337, 2020.
- [6] A. Khusnul Khotimah, “Analisis Sentimen Terhadap Kualitas Pelayanan,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 3044–3048, 2024, doi: 10.36040/jati.v8i3.9520.
- [7] R. Puspitasari, “Analisis Sentimen Opini Publik Terhadap Kebijakan Baru Skripsi Pada Media Sosial Twitter Menggunakan Metode *Naive Bayes*,” vol. 05, no. 03, pp. 37–42, 2024.
- [8] A. Syafa, O. Nurdiawan, and F. M. Basysyar, “*Naive Bayes* Meningkatkan *Model* Analisis Sentimen,” vol. 9, no. 2, pp. 171–180, 2025.
- [9] J. Ha, M. Kambe, and J. Pe, *Data Mining: Concepts and Techniques*. 2011. doi: 10.1016/C2009-0-61819-5.
- [10] S. Kongyoung, C. Macdonald, and I. Ounis, “monoQA: Multi-Task Learning of Reranking and Answer Extraction for Open-Retrieval Conversational Question Answering,” *Proc. 2022 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2022*, pp. 7207–7218, 2022, doi: 10.18653/v1/2022.emnlp-main.485.
- [11] A. Firdaus, *Analisis Sentimen Pada Aplikasi Alfagift Menggunakan Metode Naive Bayes Classifier*, vol. 13, no. 1. 2023.
- [12] I. T. Julianto, D. Kurniadi, Y. Septiana, and A. Sutedi, “Pra-Pemrosesan Teks Alternatif Menggunakan Chat Gpt Buka,” vol. 12, pp. 67–77, 2023.
- [13] G. Tamami, W. A. Triyanto, and S. Muzid, “Sentiment Analysis Mobile JKN Reviews Using SMOTE Based LSTM,” vol. 19, no. 1, pp. 13–24, 2025.
- [14] M. Hermiani *et al.*, “Analisis Ulasan Produk Menggunakan Metode *Naive Bayes Classifier*,” *Jnatia*, vol. 2, no. 1, pp. 177–184, 2023.
- [15] A. Nurian, M. S. Ma’arif, I. N. Amalia, and C. Rozikin, “Analisis Sentimen Pengguna Aplikasi Shopee Pada Situs Google Play Menggunakan *Naive Bayes Classifier*,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, 2024, doi: 10.23960/jitet.v12i1.3631.
- [16] A. Syahrir, Harlinda, and F. Umar, “Analisis Sentimen Masyarakat Terhadap Kebijakan Pemerintah Vaksinasi Booster 2 Menggunakan Metode *Naive Bayes Classifier*,” *Zo. J. Sist. Inf.*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.