

**ANALISIS SENTIMEN PENGGUNA *YOUTUBE* TERHADAP
DAMPAK *ARTIFICIAL INTELLIGENCE* MENGGUNAKAN
ALGORITMA *NAIVE BAYES***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

CERLANG REALIS WIJAYA

21.12.2130

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN PENGGUNA *YOUTUBE* TERHADAP
DAMPAK *ARTIFICIAL INTELLIGENCE* MENGGUNAKAN
ALGORITMA *NAIVE BAYES***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

CERLANG REALIS WIJAYA

21.12.2130

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA *YOUTUBE* TERHADAP DAMPAK
ARTIFICIAL INTELLIGENCE MENGGUNAKAN ALGORITMA *NAIVE*
*BAYES***

yang disusun dan diajukan oleh

CELANG REALIS WIJAYA

21.12.2130

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Agustus 2025

Dosen Pembimbing,



Norhikmah, S.Kom., M.Kom.

NIK. 190302245

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN PENGGUNA *YOUTUBE* TERHADAP DAMPAK
ARTIFICIAL INTELLIGENCE* MENGGUNAKAN ALGORITMA *NAIVE
BAYES

yang disusun dan diajukan oleh

CELANG REALIS WIJAYA

21.12.2130

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Ria Andriani, S.Kom., M.Kom.

NIK. 190302458

Majid Rahardi, S.Kom., M.Eng.

NIK. 190302393

Norhikmah, S.Kom., M.Kom.

NIK. 190302245

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Cerlang Realis Wijaya
NIM : 21.12.2130

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN PENGGUNA *YOUTUBE* TERHADAP DAMPAK
ARTIFICIAL INTELLIGENCE MENGGUNAKAN ALGORITMA *NAIVE
BAYES***

Dosen Pembimbing : Norhikmah, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 Agustus 2025

Yang Menyatakan,



Cerlang Realis Wijaya

HALAMAN PERSEMBAHAN

Dengan rasa syukur kepada Tuhan Yang Maha Esa, karya tulis ini saya persembahkan kepada keluarga yang senantiasa memberikan dukungan dan semangat selama proses penyusunan skripsi, kepada dosen pembimbing yang telah membimbing dengan penuh kesabaran, serta kepada seluruh pihak yang telah memberikan bantuan, baik secara langsung maupun tidak langsung dalam menyelesaikan tugas akhir ini.



KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan bimbingan-Nya sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik. Laporan ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi di Universitas Amikom Yogyakarta.

Penulis menyadari bahwa penyusunan tugas akhir ini tidak akan terselesaikan tanpa adanya dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr. Kusrini., M.Kom selaku Dekan Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto, M.Kom selaku Ketua Program Studi Sistem Informasi.
4. Ibu Norhikmah, S.Kom., M.Kom selaku dosen pembimbing yang telah memberikan arahan, saran, dan motivasi selama proses penyusunan tugas akhir ini.
5. Kedua orang tua, keluarga besar, dan teman teman tercinta atas semangat, doa, serta dukungan moral yang selalu diberikan.
6. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu yang turut memberikan dukungan selama penyusunan tugas akhir ini.

Akhir kata, penulis berharap tugas akhir ini dapat memberikan manfaat bagi pembaca dan menjadi kontribusi positif dalam pengembangan ilmu pengetahuan di bidang sistem informasi.

Yogyakarta, 1 Agustus 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	14

2.2.1	Google Colab	14
2.2.2	Machine Learning	14
2.2.3	Analisis Sentimen	14
2.2.4	Dampak Artificial Intelligence (AI).....	15
2.2.5	YouTube	15
2.2.6	Data Mining	16
2.2.7	Klasifikasi	16
2.2.8	Bahasa Python.....	16
2.2.9	Preprocessing	17
2.2.10	Labelling	18
2.2.11	TF-IDF	18
2.2.12	Split Data	19
2.2.13	SMOTE	19
2.2.14	Naive Bayes	20
2.2.15	Smoothing	23
2.2.16	Confusion Matrix	24
2.2.17	Streamlit	26
BAB III METODE PENELITIAN		27
3.1	Objek Penelitian.....	27
3.2	Alur Penelitian	27
3.3	Alat dan Bahan.....	32
3.3.1	Alat.....	32
3.3.2	Bahan	33
BAB IV HASIL DAN PEMBAHASAN		34
4.1	Crawling data.....	34

4.2	<i>Preprocessing</i>	36
4.3	<i>Labelling</i>	42
4.4	TF-IDF	44
4.5	Split Data	45
4.6	SMOTE	46
4.7	Implementasi <i>Naive Bayes</i>	47
4.7.1	Multinomial	48
4.7.2	Gaussian	52
4.7.3	Bernoulli	53
4.8	Evaluasi Model	55
4.9	<i>Deployment</i>	55
4.9.1	Halaman Dashboard	56
4.9.2	Halaman Prediksi Sentimen	56
4.10	Testing Aplikasi	57
BAB V	PENUTUP	59
5.1	Kesimpulan	59
5.2	Saran	59
REFERENSI		61

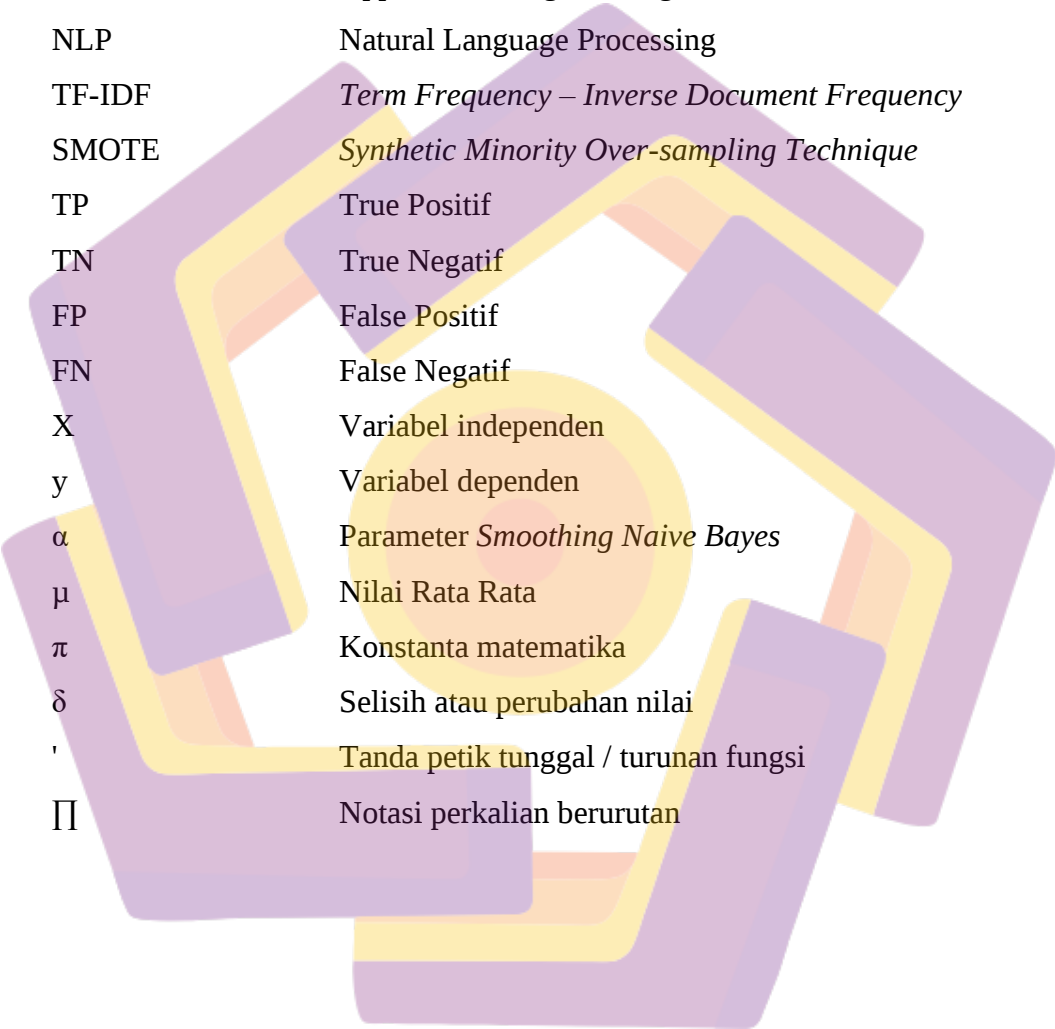
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	9
Tabel 2.2 <i>Confusion Matrix</i>	25
Tabel 4.1 Channel dan Link.....	34
Tabel 4.2 Hasil <i>Cleaning</i>	38
Tabel 4.3 Hasil <i>Case Folding</i>	38
Tabel 4.4 Hasil Normalisasi.....	39
Tabel 4.5 Hasil Tokenisasi.....	40
Tabel 4.6 Hasil <i>Stopword</i>	41
Tabel 4.7 Hasil <i>Stemming</i>	42
Tabel 4.8 Hasil <i>Labelling</i>	44
Tabel 4.9 Hasil Klasifikasi <i>MultinomialNB</i>	49
Tabel 4.10 <i>Confusion Matrix MultinomialNB</i>	49
Tabel 4.11 Hasil Klasifikasi Dengan Alpha 0.1.....	51
Tabel 4.12 Klasifikasi Dengan Alpha 0.3	51
Tabel 4.13 Hasil Klasifikasi Dengan Alpha 0.5.....	51
Tabel 4.14 <i>Confusion Matrix MultinomialNB</i>	52
Tabel 4.15 Hasil Klasifikasi <i>GaussianNB</i>	53
Tabel 4.16 <i>Confusion Matrix GaussianNB</i>	53
Tabel 4.17 Hasil Klasifikasi <i>BernoulliNB</i>	54
Tabel 4.18 <i>Confusion Matrix BernoulliNB</i>	54
Tabel 4.19 Perbandingan Hasil Akurasi <i>Naive Bayes</i>	55
Tabel 4.20 Hasil Testing	57

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	28
Gambar 4.1 <i>Crawling</i> Data 1	34
Gambar 4.2 <i>Crawling</i> Data 2	35
Gambar 4.3 Tabel Komentar.....	36
Gambar 4.4 Skrip <i>Cleaning</i>	37
Gambar 4.5 Skrip <i>Cleaning</i>	37
Gambar 4.6 Skrip <i>Case Folding</i>	38
Gambar 4.7 Skrip Normalisasi.....	39
Gambar 4.8 Skrip Tokenisasi.....	40
Gambar 4.9 Skrip <i>Stopword</i>	41
Gambar 4.10 Skrip <i>Stemming</i>	42
Gambar 4.11 Skrip <i>Labelling</i> 1.....	43
Gambar 4.12 Skrip <i>Labelling</i> 2.....	44
Gambar 4.13 Skrip TF-IDF.....	44
Gambar 4.14 Hasil TF-IDF.....	45
Gambar 4.15 Skrip Split Data.....	46
Gambar 4.16 Hasil Split Data	46
Gambar 4.17 Skrip SMOTE	46
Gambar 4.18 Hasil SMOTE.....	47
Gambar 4.19 Skrip Implementasi <i>Naive Bayes</i>	47
Gambar 4.20 Skrip Pengujian <i>MultinomialNB</i>	48
Gambar 4.21 Skrip Pengujian <i>Multinomial Alpha</i>	50
Gambar 4.22 Hasil Pengujian <i>Alpha</i>	50
Gambar 4.23 Skrip Pengujian <i>GaussianNB</i>	52
Gambar 4.24 Skrip Pengujian <i>BernoulliNB</i>	54
Gambar 4.25 Halaman Dashboard	56
Gambar 4. 26 Halaman Prediksi Sentimen	56

DAFTAR LAMBANG DAN SINGKATAN



<i>AI</i>	<i>Artificial Intelligence</i>
<i>NB</i>	<i>Naive Bayes</i>
<i>KDD</i>	<i>Knowledge Discovery in Databases</i>
<i>API</i>	<i>Application Programming Interface</i>
<i>NLP</i>	<i>Natural Language Processing</i>
<i>TF-IDF</i>	<i>Term Frequency – Inverse Document Frequency</i>
<i>SMOTE</i>	<i>Synthetic Minority Over-sampling Technique</i>
<i>TP</i>	<i>True Positif</i>
<i>TN</i>	<i>True Negatif</i>
<i>FP</i>	<i>False Positif</i>
<i>FN</i>	<i>False Negatif</i>
<i>X</i>	<i>Variabel independen</i>
<i>y</i>	<i>Variabel dependen</i>
α	<i>Parameter Smoothing Naive Bayes</i>
μ	<i>Nilai Rata Rata</i>
π	<i>Konstanta matematika</i>
δ	<i>Selisih atau perubahan nilai</i>
$'$	<i>Tanda petik tunggal / turunan fungsi</i>
$\prod$	<i>Notasi perkalian berurutan</i>

DAFTAR ISTILAH



Analisis Sentimen	Proses mengelompokkan opini menjadi positif, negatif, atau netral
<i>Crawling</i>	Proses pengambilan data dari internet secara otomatis
<i>Preprocessing</i>	Tahapan awal pembersihan dan pemrosesan teks
<i>Labelling</i>	Proses memberikan kelas pada data
TF-IDF	Bobot kata berdasarkan frekuensi lokal dan global
Split Data	Pembagian data menjadi data latih dan data uji
SMOTE	Teknik penyeimbangan data dengan membuat data sintetis
<i>Naive Bayes</i>	Algoritma klasifikasi berbasis probabilitas
<i>Multinomial</i>	<i>Naive Bayes</i> untuk data berupa jumlah kemunculan kata
<i>Bernoulli</i>	<i>Naive Bayes</i> untuk data biner
<i>Gaussian</i>	<i>Naive Bayes</i> untuk data numerik kontinu
<i>Alpha</i>	Parameter <i>Smoothing</i> untuk menghindari probabilitas nol
<i>Confusion Matrix</i>	Matriks evaluasi hasil klasifikasi
<i>Deployment</i>	Proses menerapkan model ke sistem nyata agar bisa digunakan oleh pengguna

INTISARI

Perkembangan Artificial Intelligence (AI) yang pesat memunculkan beragam respons dari masyarakat, mulai dari dukungan hingga kekhawatiran terhadap dampak negatif yang ditimbulkan. Media sosial seperti YouTube menjadi wadah utama bagi masyarakat untuk mengekspresikan opini tersebut melalui kolom komentar, sehingga analisis sentimen penting dilakukan guna memahami kecenderungan opini publik terhadap dampak AI. Penelitian ini menganalisis sentimen masyarakat terhadap dampak AI berdasarkan komentar YouTube menggunakan tiga varian algoritma *Naive Bayes*, yaitu *Multinomial*, *Gaussian*, dan *Bernoulli* serta menguji pengaruh variasi nilai parameter *Alpha* terhadap algoritma *Multinomial Naive Bayes*. Data yang digunakan telah melalui *Preprocessing* dan diklasifikasikan menggunakan metode *Naive Bayes*, dengan pengukuran performa melalui metrik akurasi, presisi, recall, dan F1-score.

Hasil pengujian menunjukkan bahwa *Multinomial Naive Bayes* memberikan performa terbaik dalam klasifikasi sentimen, dengan akurasi tertinggi sebesar 90,28%, precision sebesar 94% dan 86%, recall 87% dan 94%, serta F1-score 91% dan 90% untuk masing-masing kelas. Sementara itu, algoritma *Gaussian* dan *Bernoulli* menunjukkan performa lebih rendah, dengan akurasi masing-masing sebesar 68,06% dan 68,52%. Pengujian parameter *Alpha* pada *Multinomial Naive Bayes* menunjukkan bahwa nilai *Alpha* mempengaruhi akurasi model, di mana *Alpha* 0.1 menghasilkan akurasi 87,04%, dan *Alpha* 0.3 serta 0.5 sebesar 87,50%. Dengan demikian, *Multinomial Naive Bayes* tanpa parameter *Alpha* terbukti paling efektif untuk klasifikasi sentimen berbasis teks komentar.

Kata kunci: Analisis Sentimen, Artificial Intelligence, YouTube, Naive Bayes, Alpha.

ABSTRACT

The rapid development of Artificial Intelligence (AI) has generated diverse responses from society, ranging from support to concerns about its potential negative impacts. Social media platforms such as YouTube serve as major spaces for people to express their opinions through comments, making sentiment analysis essential to understand public perceptions of AI. This study analyzes public sentiment regarding the impact of AI based on YouTube comments using three variants of the Naive Bayes algorithm Multinomial, Gaussian, and Bernoulli and examines the effect of varying Alpha parameters on the Multinomial Naive Bayes algorithm. The collected data underwent preprocessing and was classified using the Naive Bayes method, with performance evaluated through accuracy, precision, recall, and F1-score metrics.

The results indicate that Multinomial Naive Bayes delivers the best performance in sentiment classification, achieving the highest accuracy of 90.28%, with precision scores of 94% and 86%, recall scores of 87% and 94%, and F1-scores of 91% and 90% for each respective class. In contrast, the Gaussian and Bernoulli algorithms showed lower performance, with accuracies of 68.06% and 68.52%, respectively. The Alpha parameter testing on Multinomial Naive Bayes demonstrated that the Alpha value influences the model's accuracy, where an Alpha of 0.1 yielded an accuracy of 87.04%, and Alphas of 0.3 and 0.5 both resulted in 87.50%. Thus, the Multinomial Naive Bayes algorithm without the Alpha parameter proved to be the most effective for text-based sentiment classification of YouTube comments.

Keyword: Sentiment Analysis, Artificial Intelligence, YouTube, Naive Bayes, Alpha.