

**ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP
VIDEO RENOVASI MUSIK EKA GUSTIWANA
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE
BAYES**



SKRIPSI

**Henk Natalis Aibekob
00000065012**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP
VIDEO RENOVASI MUSIK EKA GUSTIWANA
MENGUNAKAN ALGORITMA MULTINOMIAL NAIVE
BAYES**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**Henk Natalis Aibekob
00000065012**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Henk Natalis Aibekob
Nomor Induk Mahasiswa : 00000065012
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Komentar YouTube terhadap Video Renovasi Musik Eka Gustiwana Menggunakan Algoritma Multinomial Naive Bayes

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 30 Juni 2025



(Henk Natalis Aibekob)

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP VIDEO RENOVASI MUSIK EKA GUSTIWANA MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE BAYES

oleh

Nama : Henk Natalis Aibekob
NIM : 00000065012
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 17 Juli 2025

Pukul 15.00 s/d 17.00 dan dinyatakan

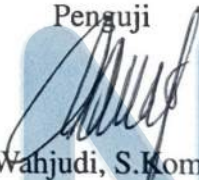
LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang


(Aditiyawan, S.Komp., M.Si.)
NIDK: 8994550022

Penguji


(Januar Wahjudi, S.Kom., M.Sc.)
NIDN: 0330017201

Pembimbing


(Alethea Suryadibrata, S.Kom., M.Eng.)
NIDN: 0322099201

Ketua Program Studi Informatika,


(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)
NIDN: 0315109103

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Henk Natalis Aibekob
NIM : 00000065012
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen Komentar YouTube
terhadap Video Renovasi Musik Eka
Gustiwana Menggunakan Algoritma
Multinomial Naive Bayes

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 30 Juni 2025

Yang menyatakan



Henk Natalis Aibekob

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji syukur atas berkat dan rahmat dari Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: **Analisis Sentimen Komentar YouTube terhadap Video Renovasi Musik Eka Gustiwana Menggunakan Algoritma Multinomial Naive Bayes**, yang dilakukan untuk memenuhi salah satu syarat guna mencapai gelar Sarjana Komputer Jurusan Informatika pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, sejak masa perkuliahan hingga penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Alethea Suryadibrata, S.Kom., M.Eng., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Bapa, Mama, Arend, Abigail dan Deslin yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan skripsi ini.

Semoga skripsi ini berguna sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 30 Juni 2025



Henk Natalis Aibekob

ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP VIDEO RENOVASI MUSIK EKA GUSTIWANA MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE BAYES

Henk Natalis Aibekob

ABSTRAK

Analisis sentimen digunakan untuk memahami persepsi atau tanggapan publik terhadap suatu topik, produk, layanan, atau peristiwa tertentu, terutama melalui data dari media sosial. Komentar pengguna di platform YouTube mencerminkan opini dan respons publik terhadap suatu konten, salah satunya video “*Renovasi Musik — TURU NIGHT*” oleh Eka Gustiwana. Penelitian ini bertujuan untuk mengklasifikasikan sentimen komentar YouTube terhadap video tersebut ke dalam kategori positif dan negatif menggunakan algoritma *Multinomial Naive Bayes*. Data yang digunakan merupakan komentar berbahasa Indonesia yang diperoleh dengan metode *crawling*, kemudian diproses melalui tahapan *text preprocessing* seperti *cleansing*, *case folding*, *tokenization*, *stopword removal*, *normalization*, dan *stemming*. Pelabelan dilakukan secara otomatis menggunakan pendekatan *lexicon-based* berdasarkan daftar kata positif dan negatif. Fitur teks dikonversi menggunakan *TF-IDF Vectorizer*. Penelitian ini juga menerapkan metode SMOTE untuk menangani ketidakseimbangan data dan membandingkan performa model dengan dan tanpa SMOTE, baik menggunakan validasi *train-test split* maupun K-Fold. Hasil pengujian menunjukkan bahwa model terbaik diperoleh pada pengujian dengan SMOTE menggunakan pembagian data latih dan uji 80:20, yang menghasilkan akurasi sebesar 85,08%, dengan nilai *precision* 85%, *recall* 85%, dan *f1-score* 85%.

Kata kunci: Analisis Sentimen, *Multinomial Naive Bayes*, Renovasi Musik, SMOTE, YouTube

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

**SENTIMENT ANALYSIS OF YOUTUBE COMMENTS ON EKA
GUSTIWANA "RENOVASI MUSIK" VIDEO USING THE MULTINOMIAL
NAIVE BAYES ALGORITHM**

Henk Natalis Aibekob

ABSTRACT

Sentiment analysis is used to understand public perception or response toward a particular topic, product, service, or event, especially through data obtained from social media. User comments on the YouTube platform reflect public opinions and responses toward content, such as the video "Renovasi Musik — TURU NIGHT" by Eka Gustiwana. This study aims to classify the sentiment of YouTube comments on the video into positive and negative categories using the Multinomial Naive Bayes algorithm. The data used are Indonesian-language comments obtained through the crawling method and processed through several text preprocessing stages, including cleansing, case folding, tokenization, stopword removal, normalization, and stemming. Sentiment labeling was performed automatically using a lexicon-based approach based on lists of positive and negative words. Text features were converted using TF-IDF Vectorizer. This study also applied the SMOTE method to address data imbalance and compared the model's performance with and without SMOTE, using both train-test split and K-Fold validation. The test results show that the best-performing model was obtained using SMOTE with an 80:20 train-test split, resulting in an accuracy of 85.08%, with a precision of 85%, recall of 85%, and f1-score of 85%.

Keywords: *Multinomial Naive Bayes, Music Renovation, Sentiment Analysis, SMOTE, YouTube*

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 YouTube	5
2.3 Text Pre-Processing	5
2.4 Labeling	6
2.5 SMOTE	7
2.6 Multinomial Naive Bayes	8
2.7 TermFrequency-Inverse Document Frequency (TF-IDF)	9
2.8 Confusion Matrix	10
2.9 K-fold Cross-Validation	12
BAB 3 METODOLOGI PENELITIAN	13
3.1 Gambaran Umum Alur Penelitian	13
3.2 Metodologi Penelitian	14
3.2.1 Studi Literatur	14
3.2.2 Crawling Data	14
3.2.3 Text-Preprocessing	14
3.2.4 Labeling	16
3.2.5 TF-IDF	16
3.2.6 Multinomial Naive Bayes	17
3.2.7 Evaluation and Testing	20
BAB 4 HASIL DAN DISKUSI	21
4.1 Spesifikasi Sistem	21
4.2 Implementasi Sistem	21
4.2.1 Potongan Kode Crawling Data	21
4.2.2 Potongan Kode Text-Preprocessing	23
4.2.3 Potongan Kode Labeling	27
4.2.4 Potongan Kode Pembobotan TF-IDF Vectorizer	29

4.2.5	Potongan Kode Pembagian Data Training dan Testing	30
4.2.6	Potongan Kode K-Fold Data Training dan Evaluasi Model .	30
4.3	Uji Coba dan Evaluasi	31
4.3.1	Pengujian Tanpa Menggunakan SMOTE Pada Data Train . .	31
4.3.2	Pengujian Menggunakan SMOTE Pada Data Train	33
4.3.3	Pembuktian Validitas Skema 80:20	34
BAB 5	SIMPULAN DAN SARAN	36
5.1	Simpulan	36
5.2	Saran	36
DAFTAR PUSTAKA		37



DAFTAR TABEL

Tabel 2.1	<i>Confusion Matrix</i>	11
Tabel 4.1	Hasil <i>Cleansing</i>	24
Tabel 4.2	Hasil <i>Case Folding</i>	24
Tabel 4.3	Hasil <i>Tokenization</i>	25
Tabel 4.4	Hasil <i>Stopword Removal</i>	25
Tabel 4.5	Hasil <i>Normalization</i>	26
Tabel 4.6	Hasil <i>Stemming</i>	27



DAFTAR GAMBAR

Gambar 3.1	<i>Flowchart utama</i>	13
Gambar 3.2	<i>Flowchart Text Pre-Processing</i>	15
Gambar 3.3	<i>Flowchart TF-IDF</i>	17
Gambar 3.4	<i>Flowchart Multinomial Naive Bayes</i>	19
Gambar 4.1	Hasil <i>Labeling</i>	29
Gambar 4.2	Hasil Pengujian Tanpa Menggunakan SMOTE Pada Data Train	32
Gambar 4.3	<i>Confusion Matrix</i> Pengujian Tanpa Menggunakan SMOTE Pada Data Train	32
Gambar 4.4	Hasil Pengujian Menggunakan SMOTE Pada Data Train	33
Gambar 4.5	<i>Confusion Matrix</i> Pengujian Menggunakan SMOTE Pada Data Train	34
Gambar 4.6	Hasil K-Fold	35



DAFTAR KODE

Kode 4.1	Potongan Kode <i>Crawling Data</i>	22
Kode 4.2	Potongan Kode <i>Text-Preprocessing</i>	23
Kode 4.3	Lanjutan 1 Potongan Kode <i>Text-Preprocessing</i>	26
Kode 4.4	Lanjutan 2 Potongan Kode <i>Text-Preprocessing</i>	27
Kode 4.5	Potongan Kode <i>Labeling</i>	28
Kode 4.6	Potongan Kode TF-IDF	29
Kode 4.7	Potongan Kode Pembagian Data <i>Training</i> dan <i>Testing</i>	30
Kode 4.8	Potongan Kode K-Fold Data <i>Training</i> dan Evaluasi Model	31



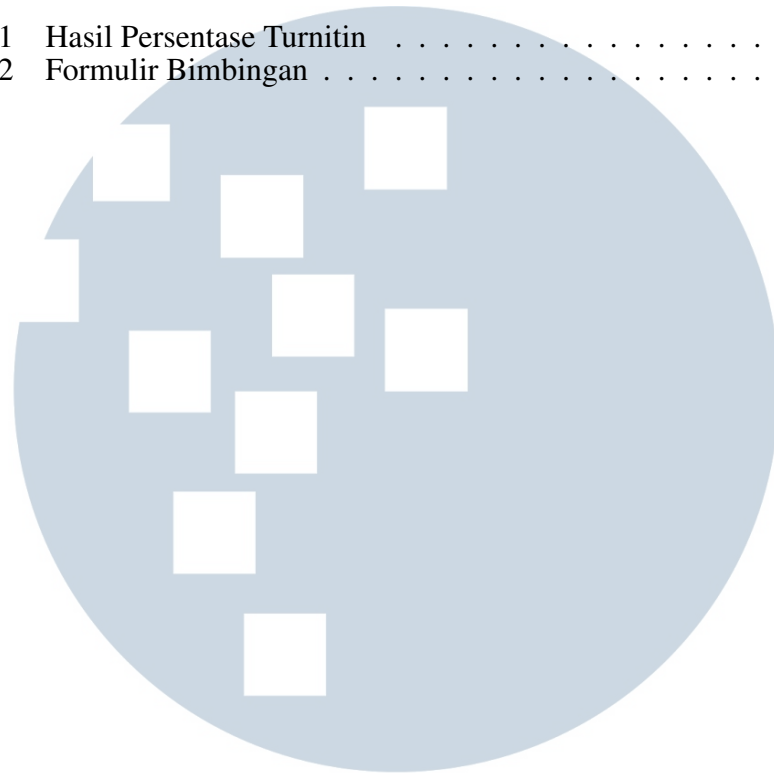
DAFTAR RUMUS

Rumus 2.1	<i>Rumus Dasar Naive Bayes</i>	8
Rumus 2.2	<i>Multinomial Naive Bayes</i>	8
Rumus 2.3	<i>Prior Probability</i>	8
Rumus 2.4	<i>Conditional Probability</i>	9
Rumus 2.5	<i>Laplace Smoothing</i>	9
Rumus 2.6	<i>Term Frequency (TF)</i>	9
Rumus 2.7	<i>Inverse Document Frequency (IDF)</i>	10
Rumus 2.8	<i>TF-IDF</i>	10
Rumus 2.9	<i>Precision</i>	11
Rumus 2.10	<i>Recall</i>	11
Rumus 2.11	<i>Accuracy</i>	11
Rumus 2.12	<i>F1 Score</i>	11



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	39
Lampiran 2	Formulir Bimbingan	40



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA