

**PENGEMBANGAN SISTEM ANALISIS SENTIMEN
KOMENTAR VIDEO YOUTUBE DENGAN METODE
PROTOTYPING UNTUK MEMBANTU PEMILIHAN KONTEN
BERKUALITAS**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

GABRIEL GALANG RESTU AJI

21.61.0222

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**PENGEMBANGAN SISTEM ANALISIS SENTIMEN
KOMENTAR VIDEO YOUTUBE DENGAN METODE
PROTOTYPING UNTUK MEMBANTU PEMILIHAN KONTEN
BERKUALITAS**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

GABRIEL GALANG RESTU AJI

21.61.0222

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**PENGEMBANGAN SISTEM ANALISIS SENTIMEN KOMENTAR
VIDEO YOUTUBE DENGAN METODE PROTOTYPING UNTUK
MEMBANTU PEMILIHAN KONTEN BERKUALITAS**

yang disusun dan diajukan oleh

Gabriel Galang Restu Aji

21.61.0222

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 11 Desember 2024

Dosen Pembimbing,


Alind Yugin, S.Kom., M.Kom.

NIK. 190302255

HALAMAN PENGESAHAN

SKRIPSI

**PENGEMBANGAN SISTEM ANALISIS SENTIMEN KOMENTAR VIDEO
YOUTUBE DENGAN METODE PROTOTYPING UNTUK MEMBANTU
PEMILIHAN KONTEN BERKUALITAS**

yang disusun dan diajukan oleh

Gabriel Galang Restu Aji

21.61.0222

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theophilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Ferian Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

Ainul Yaqin, S.Kom., M.Kom.
NIK. 190302255



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom
NIK. 190302106

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Gabriel Galang Restu Aji
NIM : 21.61.0222

Menyatakan bahwa Skripsi dengan judul berikut:

PENGEMBANGAN SISTEM ANALISIS SENTIMEN KOMENTAR VIDEO YOUTUBE DENGAN METODE PROTOTYPING UNTUK MEMBANTU PEMILIHAN KONTEN BERKUALITAS

Dosen Pembimbing : Ainul Yaqin, 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, 23 Juli 2025

Yang Menyatakan,



Gabriel Galang Restu Aji

HALAMAN PERSEMBAHAN

Dengan mengucapkan syukur Alhamdulillah, segala puji bagi Allah SWT atas rahmat dan keberkahan-Nya, sehingga skripsi ini dapat diselesaikan dengan baik. Skripsi ini saya persembahkan kepada:

1. Kedua orang tua dan seluruh keluarga yang senantiasa memberikan doa, dukungan, semangat, serta motivasi tanpa henti dalam setiap langkah perjalanan akademik saya.
2. Bapak Ferian Fauzi Abdulloh, S.Kom., M.Kom., selaku dosen wali, serta Bapak Ainul Yaqin, S.Kom., M.Kom., selaku dosen pembimbing, atas kesabaran, bimbingan, serta ilmu yang telah diberikan selama proses penyusunan skripsi ini.
3. Veranica, yang telah menjadi sumber dukungan, semangat, serta motivasi selama masa perkuliahan dan penyusunan skripsi ini.
4. Sahabat-sahabat terbaik saya: Mutamammil Ula Herwanto, Rizqy Resha Prameswara, Farhan Hidayatul Haq, Septa Firdaus Adianta, Ririn Nurdianti, dan Dhanu Ahsan Fatih, yang selalu menemani, mendukung, serta memberikan semangat sepanjang perjalanan akademik ini.
5. Seluruh rekan 21 BCI 01, atas kebersamaan, dukungan, dan pengalaman berharga yang telah dibagikan selama masa perkuliahan.

Semoga segala kebaikan dan dukungan yang telah diberikan mendapatkan balasan yang berlipat ganda.

KATA PENGANTAR

Puji dan syukur saya panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya, sehingga saya dapat menyelesaikan skripsi ini sebagai salah satu syarat akademik untuk memperoleh gelar Sarjana Komputer di Universitas AMIKOM Yogyakarta. Penulisan skripsi ini tidak terlepas dari dukungan, bimbingan, serta bantuan dari berbagai pihak, yang dengan tulus telah memberikan ilmu, motivasi, dan arahan selama proses penyusunannya. Oleh karena itu, dengan segala kerendahan hati, saya ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah berkontribusi dalam penyelesaian skripsi ini.

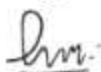
Skripsi ini juga menjadi bukti bahwa penulis telah menyelesaikan seluruh tahapan perkuliahan dalam program Strata-1. Dalam kesempatan ini, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Ainul Yaqin, S.Kom., M.Kom., selaku dosen pembimbing, atas bimbingan, arahan, serta dukungan yang telah diberikan selama proses penyusunan skripsi ini.
4. Seluruh dosen Universitas AMIKOM Yogyakarta, yang telah berbagi ilmu dan wawasan yang berharga selama masa perkuliahan.
5. Rekan-rekan mahasiswa, khususnya keluarga besar 21 BCI 01, atas dukungan, kebersamaan, dan pengalaman berharga yang telah diberikan selama menjalani perkuliahan.
6. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah berkontribusi dalam penyelesaian skripsi ini.

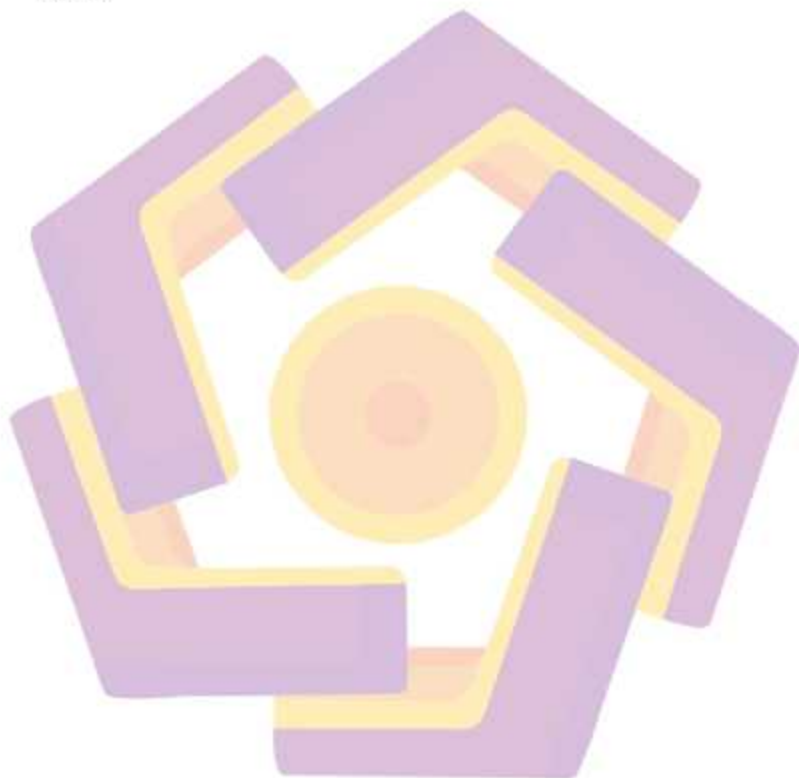
Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik serta saran yang

membangun untuk penyempurnaan lebih lanjut. Besar harapan penulis agar skripsi ini dapat memberikan manfaat bagi berbagai pihak yang membacanya.

Yogyakarta, 19 Juli 2025



Penulis



DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvi
DAFTAR LAMBANG DAN SINGKATAN	xvii
DAFTAR ISTILAH	xviii
INTISARI.....	xix
ABSTRACT.....	xx
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur	7
2.2 Dasar Teori.....	20

2.2.1.	Sentimen Analisis	20
2.2.2.	Beberapa Kegunaan Sentimen Analisis	22
2.2.3.	Sentimen Analisis Dalam Media Sosial	23
2.2.4.	Pemilihan Model Sentimen Analisis	24
2.2.5.	Perbandingan Model Sentimen Analisis	26
2.2.6.	Perbandingan Dengan Penelitian Lain	27
2.2.7.	Pengertian Metode Prototyping	28
2.2.8.	Kelebihan Metode Prototyping	30
2.2.9.	Kekurangan Metode Prototyping	31
2.2.10.	Kecocokan Metode Prototyping	31
2.2.11.	<i>Preprocessing</i> dan Langkahnya	32
2.2.12.	Evaluasi Performa Klasifikasi	38
2.2.13.	Perhitungan Polaritas (Metode Leksikon)	41
2.2.14.	Komparasi Metode Sentimen	46
BAB III METODE PENELITIAN		50
3.1	Objek Penelitian	50
3.2	Alur Penelitian	50
3.2.1.	Akuisisi Data	52
3.2.2.	Data Cleaning	52
3.2.3.	<i>Case folding</i>	53
3.2.4.	Normalisasi atau <i>Slang replacement</i>	53
3.2.5.	<i>Stopword Removal</i>	54
3.2.6.	Pelabelan Leksikon	54
3.2.7.	Feature Extraction	59
3.2.8.	<i>Multinomial Naïve Bayes</i>	64
3.2.9.	<i>Linear SVC</i>	67

3.2.10.	Contoh Evaluasi <i>F1-Score</i>	70
3.2.11.	Perumusan Kebutuhan Awal.....	71
3.2.12.	Pembuatan Prototipe	72
3.2.13.	Pengujian dan <i>Deployment</i> Prototipe.....	73
3.2.14.	Umpan Balik	73
3.2.15.	<i>Final Deployment</i>	73
3.3	Alat dan Bahan.....	74
3.3.1.	Data Penelitian	74
3.3.2.	Alat/Instrumen Penelitian.....	74
BAB IV HASIL DAN PEMBAHASAN		77
4.1.	Hasil Akuisisi Data	77
4.2.	Hasil Data Cleaning	78
4.3.	Hasil <i>Case folding</i>	82
4.4.	Hasil Normalization	82
4.5.	Hasil Stopword Removal	84
4.6.	Hasil Pelabelan Leksikon.....	87
4.7.	Hasil <i>Multinomial Naïve Bayes</i> dengan BoW.....	91
4.8.	Hasil <i>Linear SVC</i> dengan TF-IDF.....	93
4.9.	Evaluasi Metode Sentimen.....	95
4.10.	Hasil Perumusan Kebutuhan Awal	98
4.11.	Hasil Pembuatan Prototipe 1	99
a.	Hasil <i>Backend</i> Prototipe 1	99
b.	Hasil <i>Frontend</i> Prototipe 1.....	103
4.12.	Hasil Pengujian dan <i>Deployment</i> Prototipe 1.....	109
4.13.	Hasil Umpan Balik 1	112
4.14.	Hasil Perumusan Kebutuhan Revisi.....	115

a. Tampilan Lebih Menarik.....	115
b. Tampilan Lebih Mudah Dipahami.....	116
c. Penambahan Fitur.....	116
d. Kecepatan Respons	117
4.15. Hasil Pembuatan Prototipe 2.....	117
a. Hasil <i>Backend</i> Prototipe 2.....	117
b. Hasil <i>Frontend</i> Prototipe 2.....	125
4.16. Hasil Pengujian dan <i>Deployment</i> Prototipe 2.....	136
4.17. Hasil <i>Umpan Balik</i> Prototipe 2	140
4.18. Hasil <i>Final Deployment</i>	144
BAB V PENUTUP.....	155
5.1 Kesimpulan	155
5.2 Saran.....	156
REFERENSI	157
LAMPIRAN.....	163

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian.....	8
Tabel 3. 1 Contoh Data Cleaning.....	52
Tabel 3. 2 Contoh Case folding.....	53
Tabel 3. 3 Normalisasi.....	53
Tabel 3. 4 Stopword Removal.....	54
Tabel 3. 5 Kamus Skor AFINN.....	54
Tabel 3. 6 Label AFINN.....	56
Tabel 3. 7 Kamus Skor SentiStrength.....	56
Tabel 3. 8 Label SentiStrength.....	58
Tabel 3. 9 Keterangan TextBlob.....	58
Tabel 3. 10 Keterangan VADER.....	59
Tabel 3. 11 Contoh Hasil Pelabelan Leksikon.....	59
Tabel 3. 12 Representasi vektor BoW.....	60
Tabel 3. 13 Perhitungan TF.....	61
Tabel 3. 14 Perhitungan IDF.....	61
Tabel 3. 15 TF-IDF Contoh 1.....	62
Tabel 3. 16 TF-IDF Contoh 2.....	62
Tabel 3. 17 TF-IDF Contoh 3.....	62
Tabel 3. 18 TF-IDF Contoh 4.....	63
Tabel 3. 19 TF-IDF Contoh 5.....	63
Tabel 3. 20 Contoh Dataset BoW Multinomial NB.....	64
Tabel 3. 21 Data Latih MNB.....	64
Tabel 3. 22 Data Uji MNB.....	65
Tabel 3. 23 Prior MNB.....	65
Tabel 3. 24 Contoh Prediksi MNB.....	67
Tabel 3. 25 Contoh Dataset LinearSVC.....	68
Tabel 3. 26 Data Latih LinearSVC.....	68
Tabel 3. 27 Data Uji LinearSVC.....	68
Tabel 3. 28 Fungsi Objektif LinearSVC.....	69

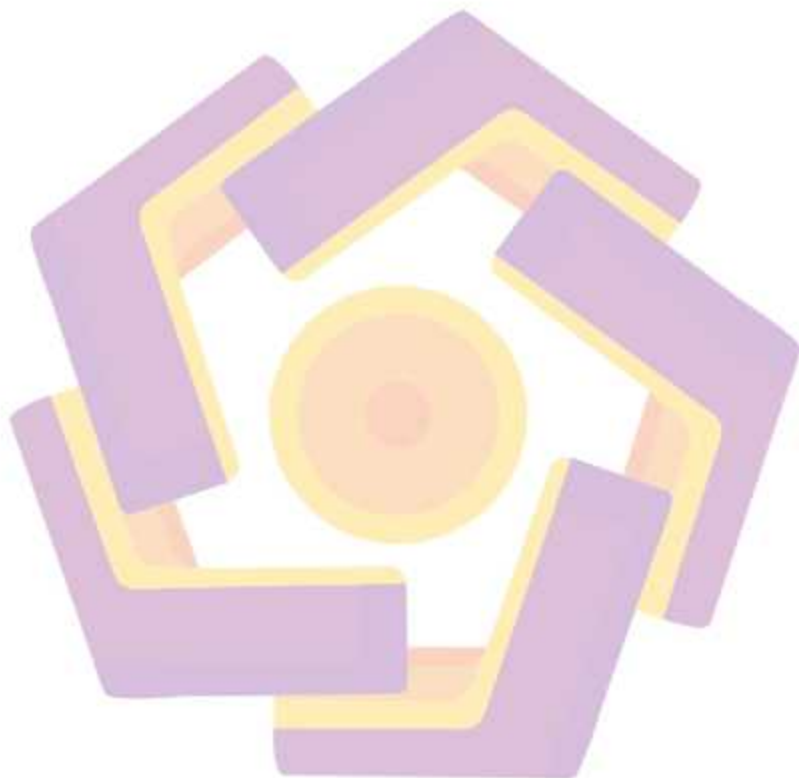
Tabel 3. 29 Hasil LinearSVC	70
Tabel 3. 30 Contoh Prediksi LinearSVC.....	70
Tabel 3. 31 Contoh Evaluasi Hasil Prediksi.....	70
Tabel 4. 1 Hasil Skor Model ML	96
Tabel 4. 2 Hasil F1-Score Keseluruhan	98
Tabel 4. 3 API Prototipe 1.....	99
Tabel 4. 4 Crawling Komentar Prototipe 1	100
Tabel 4. 5 Data Cleaning Prototipe 1	101
Tabel 4. 6 Implementasi Model Prototipe 1	102
Tabel 4. 7 HTML Prototipe 1	103
Tabel 4. 8 CSS Prototipe 1	104
Tabel 4. 9 API JS Prototipe 1	107
Tabel 4. 10 Script JS Prototipe 1	109
Tabel 4. 11 API Prototipe 2.....	118
Tabel 4. 12 Crawling Komentar Prototipe 2	120
Tabel 4. 13 Pencarian dengan Keyword Prototipe 2.....	121
Tabel 4. 14 Data Cleaning Prototipe 2	122
Tabel 4. 15 Implementasi Model Prototipe 2.....	124
Tabel 4. 16 Perhitungan Persentase Label Prototipe 2	125
Tabel 4. 17 HTML Prototipe 2.....	125
Tabel 4. 18 CSS Prototipe 2.....	127
Tabel 4. 19 API JS Prototipe 2.....	132
Tabel 4. 20 Script JS Prototipe 2.....	135
Tabel 4. 21 Kritik dan Saran	143
Tabel 4. 22 Konfigurasi Nginx (http conf).....	147
Tabel 4. 23 Install Python	149
Tabel 4. 24 Aktivasi VENV dan Dependensi	150
Tabel 4. 25 Membuat Service Process	151
Tabel 4. 26 Menjalankan Service API	152

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Prototyping	29
Gambar 3. 1 Alur Penelitian.....	51
Gambar 4. 1 Download Data.....	77
Gambar 4. 2 Seleksi Kolom	78
Gambar 4. 3 Download Data.....	78
Gambar 4. 4 Data Sebelum Cleaning.....	79
Gambar 4. 5 Penghapusan Duplikat.....	79
Gambar 4. 6 Penghapusan Komentar Kosong	80
Gambar 4. 7 Menghapus Karakter Khusus	80
Gambar 4. 8 Menghapus Unsur Tautan	81
Gambar 4. 9 Penghapusan Simbol dan Tidak Emoji	81
Gambar 4. 10 Case folding.....	82
Gambar 4. 11 Mengunduh Slang Dict.....	83
Gambar 4. 12 Pembuatan Kamus Slang.....	83
Gambar 4. 13 Fungsi Penggantian Slang dan Penerapan.....	84
Gambar 4. 14 Download Stopword.....	84
Gambar 4. 15 Membuat Custom Stopwords	85
Gambar 4. 16 Fungsi Penghapusan Stopword	85
Gambar 4. 17 Penerapan Stopword.....	85
Gambar 4. 18 Penghapusan Ulang.....	86
Gambar 4. 19 Download Dataset Bersih.....	87
Gambar 4. 20 Inisialisasi Leksikon	87
Gambar 4. 21 Pelabelan TextBlob dan VADER.....	88
Gambar 4. 22 Proses Klasifikasi Leksikon	89
Gambar 4. 23 Hasil Klasifikasi Leksikon	89
Gambar 4. 24 Perbandingan Label Asli dengan TextBlob	90
Gambar 4. 25 Perbandingan Label Asli dengan VADER.....	90
Gambar 4. 26 Data Splitting MNB	91

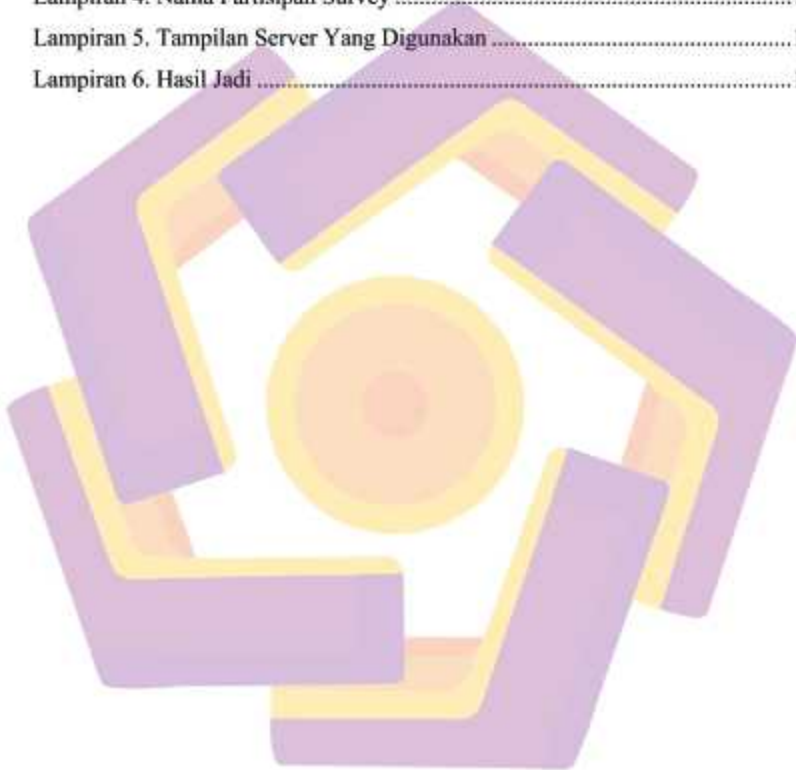
Gambar 4. 27 Ekstraksi Fitur BoW dan Training MNB	92
Gambar 4. 28 Hasil MNB	92
Gambar 4. 29 Data Splitting LinearSVC	94
Gambar 4. 30 Ekstraksi Fitur TF-IDF dan Training LinearSVC.....	94
Gambar 4. 31 Hasil LinearSVC	95
Gambar 4. 32 Hasil F1-Score TextBlob.....	97
Gambar 4. 33 Hasil F1-Score VADER	97
Gambar 4. 34 Wireframe Prototipe	98
Gambar 4. 35 Hasil Hosting Lokal	110
Gambar 4. 36 Hasil Performa NB.....	111
Gambar 4. 37 Hasil Performa SVM.....	111
Gambar 4. 38 Judul Survey Prototipe 1	113
Gambar 4. 39 Survey Prototipe 1	113
Gambar 4. 40 Responden Survey Prototipe 1	114
Gambar 4. 41 Hasil Survey Prototipe 1	114
Gambar 4. 42 Hosting Local Prototipe 2	136
Gambar 4. 43 Animasi Loading Prototipe 2	137
Gambar 4. 44 Hasil Analisa Prototipe 2	138
Gambar 4. 45 Performa NB Prototipe 2.....	139
Gambar 4. 46 Performa SVM Prototipe 2.....	139
Gambar 4. 47 Performa Textblob Prototipe 2	139
Gambar 4. 48 Performa VADER Prototipe 2.....	140
Gambar 4. 49 Judul Survey Prototipe 2	141
Gambar 4. 50 Survey Prototipe 2.....	141
Gambar 4. 51 Jumlah Responden Survey Prototipe 2.....	142
Gambar 4. 52 Hasil Survey Prototipe 2	142
Gambar 4. 53 Hasil Umpan Balik Perlunya Iterasi.....	143
Gambar 4. 54 Konfigurasi Docker Nginxwebui	144
Gambar 4. 55 Volumes Nginxwebui.....	145
Gambar 4. 56 Konfigurasi Docker PHP-fpm.....	146
Gambar 4. 57 Volumes PHP-fpm	146
Gambar 4. 58 Flask API Berjalan Lancar	152

Gambar 4. 59 Proxy Prototipe 1.....	153
Gambar 4. 60 Proxy Final Deployment.....	153
Gambar 4. 61 Hasil Hosting Dedicated Server	153



DAFTAR LAMPIRAN

Lampiran 1. Dataset Training dan Testing.....	163
Lampiran 2. Kode Sumber ipynb.....	163
Lampiran 3. Kode Sumber Sistem.....	163
Lampiran 4. Nama Partisipan Survey.....	163
Lampiran 5. Tampilan Server Yang Digunakan.....	163
Lampiran 6. Hasil Jadi.....	163



DAFTAR LAMBANG DAN SINGKATAN



API	Application Programming Interface
BoW	Bag of Words
CSS	Cascading Style Sheets
CSV	Comma-Separated Values
DL	Deep Learning
HTML	Hypertext Markup Language
JS	JavaScript
MNB	Multinomial Naïve Bayes
ML	Machine Learning
NLP	Natural Language Processing
pip	Pip install packages
Py	Python
SU	Super User
SVC	Support Vector Classification
SVM	Support Vector Machines
TF-IDF	Term Frequency-Inverse Document Frequency
UI	User Interface
UX	User Experience
VADER	Valence Aware Dictionary and sEntiment Reasoner

DAFTAR ISTILAH

Analisis Sentimen	Proses mengelompokkan opini menjadi label
Bag of Words	Representasi teks dalam kumpulan kata tanpa urutan
CasaOS	Sistem operasi berbasis web untuk mengelola server
Komentar YouTube	Tanggapan teks pengguna pada video YouTube
F1-Score	Ukuran akurasi model berdasar precision dan recall
Flask	Framework Python untuk web ringan
Lexicon-Based	Pendekatan analisis sentimen dengan kamus kata
Linear SVC	Model klasifikasi dengan garis batas linear
Multinomial Naïve Bayes	Algoritma klasifikasi teks berbasis probabilitas
Normalisasi	Mengubah kata tidak baku menjadi bentuk standar
Preprocessing	Menyiapkan teks sebelum analisis
Prototyping	Metode membuat sistem bertahap dengan iterasi
Stopword Removal	Penghapusan kata umum yang tidak penting
TF-IDF	Teknik bobot kata berdasar frekuensi dan keunikan
Tokenisasi	Pemisahan teks menjadi kata-kata (token)
WebSocket	Protokol komunikasi dua arah secara real-time
YouTube API	Alat pengambil data dari YouTube secara otomatis

INTISARI

YouTube merupakan platform video yang sangat populer dan digunakan oleh jutaan orang setiap hari sebagai sumber hiburan, informasi, dan edukasi. Namun, di balik kemudahannya, pengguna kini mengalami kesulitan dalam menilai kualitas suatu konten karena hilangnya indikator jumlah *dislike* serta meningkatnya konten buatan AI atau *content farm* yang sering kali menurunkan kualitas informasi. Pengguna harus membaca komentar secara manual untuk memahami respons audiens, yang tentu tidak efisien. Untuk menjawab masalah ini, penelitian ini mengembangkan sistem analisis sentimen komentar YouTube berbasis web dengan menggunakan metode prototyping. Sistem dirancang melalui beberapa tahapan: akuisisi komentar melalui YouTube API, *preprocessing* data teks, pelabelan sentimen menggunakan metode leksikon, pelatihan model menggunakan algoritma Multinomial Naïve Bayes dan Linear SVC, serta integrasi sistem backend dan *frontend* berbasis Flask. Hasil evaluasi menunjukkan bahwa model Linear SVC memiliki *F1-Score* tertinggi sebesar 0,7005, diikuti oleh Multinomial Naïve Bayes sebesar 0,6522, sedangkan metode leksikon seperti VADER dan TextBlob memiliki performa lebih rendah namun tetap berguna dalam analisis multibahasa. Sistem berhasil diimplementasikan dalam bentuk website yang dapat menampilkan ringkasan sentimen komentar hanya dengan memasukkan tautan atau kata kunci video. Penelitian ini bermanfaat bagi pengguna umum, kreator konten, serta peneliti di bidang pemrosesan bahasa alami, dan masih dapat dikembangkan lebih lanjut melalui penambahan model, fitur deteksi komentar bot atau spam, serta perluasan dukungan bahasa.

Kata kunci: analisis sentimen, YouTube, komentar, Naïve Bayes, Linear SVC.

ABSTRACT

YouTube is a highly popular video platform used by millions of people daily as a source of entertainment, information, and education. However, despite its convenience, users now face difficulties in evaluating content quality due to the removal of the dislike count and the increasing presence of AI-generated or content farm videos, which often degrade information quality. Users are required to manually read comments to understand audience responses, which is inefficient. To address this issue, this research developed a web-based sentiment analysis system for YouTube comments using the prototyping method. The system was built through several stages: comment acquisition via the YouTube API, text data preprocessing, sentiment labeling using lexicon-based methods, model training with Multinomial Naïve Bayes and Linear SVC algorithms, and integration of backend and frontend using the Flask framework. Evaluation results showed that the Linear SVC model achieved the highest F1-Score of 0.7063, followed by Multinomial Naïve Bayes with a score of 0.6555. Lexicon-based methods such as VADER and TextBlob produced lower F1-Scores but remained useful for multilingual sentiment analysis. The system was successfully implemented in the form of a website that can display sentiment summaries of comments by simply inputting a video link or topic. This research is useful for general users, content creators, and NLP researchers, and it can be further developed by adding more models, spam or bot detection features, and broader language support.

Keywords: sentiment analysis, YouTube, comments, Naïve Bayes, Linear SVC.