

**ANALISIS SENTIMEN TERHADAP ESPORT PUBG MOBILE
PMCO GLOBAL FINAL 2019 FINAL DENGAN METODE NAIVE
BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh
MUHAMMAD HILLAL ALHAMDY
22.11.4638

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**ANALISIS SENTIMEN TERHADAP ESPORT PUBG MOBILE PMCO
GLOBAL FINAL 2019 FINAL DENGAN METODE NAIVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh
MUHAMMAD HILLAL ALHAMDI
22.11.4638

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP ESPORT PUBG MOBILE PMCO
GLOBAL 2019 FINAL DENGAN METODE NAIVE BAYES**

yang disusun dan diajukan oleh
diajukan oleh

Muhammad Hillal Alhamdi

22.11.4638

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Januari 2026

Dosen Pembimbing,



Dina Maulina S.Kom., M.Kom

NIK. 190302250

HALAMAN PENGESAHAN

SKRIPSI
ANALISIS SENTIMEN TERHADAP ESPORT PUBG MOBILE PMCO
GLOBAL 2019 FINAL DENGAN METODE NAIVE BAYES

yang disusun dan diajukan oleh

Muhammad Hillal Alhamdi 22.11.4638

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

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

Haryoko, S.Kom., M.Cs.,
NIK. 190302286

Dina Maulina, S.Kom.,
M.Kom.,
NIK. 190302250

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Hillal Alhamdi
NIM : 22.11.4638

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN TERHADAP ESPORT PUBG MOBILE PMCO GLOBAL 2019 FINAL DENGAN METODE NAIVE BAYES

Dosen Pembimbing : Dina Maulina 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, 21 Januari 2025

Yang Menyatakan,



Muhammad Hillal Alhamdi

HALAMAN PERSEMBAHAN

Dengan menyebut nama Allah SWT Yang Maha Pengasih lagi Maha Penayayang, segala puji dan syukur senantiasa tercurah atas segala rahmat dan akrunia yang telah dilimpahkan, sehingga Skripsi ini dapat diselesaikan dengan baik.

Karya ini dengan segala kerendahan hati dan rasa syukur yang mendalam, penulis persembahkan kepada :

1. Allah SWT, atas segala petunjuk, kekuatan, dan kemudahan yang tak ada batas dalam setiap proses penulisan skripsi ini.
2. Kedua Orang Tua tercinta, Kakak, dan Adik yang penulis sayangi yang telah memberikan dukungan, doa yang tak pernah putus, memberi semangat setiap harinya, selalu mengingatkan pada jalan yang benar, serta menjadikan motivasi dan inspirasi penulis agar tidak putus asa.
3. Kepada teman-teman dekat penulis yang selalu mendengarkan cerita penulis, selalu memberikan dukungan mental, memberikan semangat tawa selama proses penulisan skripsi ini.
4. Almamater tercinta, UNIVERSITAS AMIKOM Yogyakarta

Penelitian ini diharapkan dapat membawa manfaat, memperluas pengetahuan baik bagi penulis maupun bagi pembaca, serta menjadi kontribusi kecil dalam pengembangan ilmu pengetahuan.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga skripsi berjudul "Analisis Sentimen terhadap E-Sport PUBG Mobile PMCO Global 2019 Final dengan Metode Naive Bayes" dapat diselesaikan sebagai salah satu syarat memperoleh gelar Sarjana di Universitas AMIKOM Yogyakarta.

Dalam proses penyusunan skripsi ini, penulis banyak memperoleh bantuan, dukungan, serta arahan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Mohammad Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Eli Pujastuti, M.Kom., selaku Ketua Program Studi SI Informatika Universitas AMIKOM Yogyakarta.
3. Dina Maulina, M.Kom., selaku Dosen Pembimbing, atas bimbingan, arahan, serta masukan yang diberikan selama proses penyusunan skripsi ini.
4. Seluruh dosen Universitas AMIKOM Yogyakarta, atas ilmu pengetahuan dan pengalaman yang telah diberikan kepada penulis.
5. Orang tua dan keluarga penulis, atas doa, dukungan, dan motivasi yang senantiasa diberikan.
6. Teman-teman penulis, atas dukungan, semangat, dan kebersamaan selama proses perkuliahan dan penyusunan skripsi.

Penulis menyadari keterbatasan skripsi ini dan mengharapkan kritik serta saran yang membangun agar dapat memberikan manfaat bagi pengembangan ilmu pengetahuan.

Yogyakarta, 21 Januari 202
Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvi
DAFTAR LAMBANG DAN SINGKATAN	xvii
DAFTAR ISTILAH	xviii
INTISARI	xxi
ABSTRACT	xxii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1. Manfaat Teoritis	3
2. Manfaat Praktis	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	17
2.2.1 Analisis Sentimen	17
2.2.2 Text Mining dan Natural Language Processing (NLP)	17

2.2.3	Cleaning & Filtering.....	18
2.2.5	Algoritma Naïve Bayes.....	20
2.2.6	Algoritma Support Vector Machine (SVM).....	21
2.2.7	Confusion Matrix.....	21
BAB III.....		24
METODE PENELITIAN		24
3.1	Objek Penelitian	24
1.	Fase saat fenomena berlangsung (2019)	24
2.	Fase setelah fenomena berakhir (2021-2025)	24
3.2	Alur Penelitian.....	26
1.	Web Scraping	27
2.	Filtering & Cleaning.....	27
A.	Filtering	27
B.	Cleaning	27
3.	Pembagian Data (Split Data).....	28
4.	Tahap Training dan Testing	28
A.	Jalur Training (Pelatihan).....	28
B.	Jalur Testing (Penguji).....	29
5.	Klasifikasi Algoritma (Labeling).....	29
6.	Hasil Klasifikasi.....	29
7.	Analisis Hasil Klasifikasi.....	30
	Kesimpulan Perbandingan Statistik Dataset Awal dan Dataset Updated	32
3.3	Alat dan Bahan.....	34
3.3.1	Data Penelitian	34
3.3.2	Alat dan Infrastruktur Penelitian	35
1.	Perangkat Keras.....	35

2.	Perangkat Lunak.....	35
3.	Perangkat Pendukung.....	37
3.4	Metode Pengumpulan Data.....	37
3.4.1.	Identifikasi Sumber Data.....	37
3.4.2.	Teknik Web Scraping.....	37
3.4.3.	Pembaruan Rentang Data (Update Dataset).....	39
3.4.4.	Penyaringan Bahasa (Language Filtering).....	39
3.4.5.	Pembersihan Komentar (Cleaning Data).....	41
3.5	Preprocessing Teks.....	42
3.5.1	Case Folding.....	42
3.5.7	Normalisasi Slang.....	49
3.6	Pembobotan TF-IDF.....	49
3.7	Metode Klasifikasi.....	50
3.7.1	Naive Bayes Classifier (NBC) — Metode Utama.....	50
3.7.2	Support Vector Machine (SVM) — Metode Pembanding.....	50
BAB IV		51
4.1	Implementasi.....	51
4.2	Pengambilan Data (Scraping YouTube).....	51
4.2.1	Mengakses API YouTube.....	51
4.2.2	Penyimpanan Data.....	53
4.3	Pre-Processing.....	53
4.3.1	Membaca Dataset.....	53
4.3.2	Filtering Bahasa.....	53
4.3.3	Cleaning Data.....	54
4.3.4	Case Folding.....	55

4.3.5	Tokenizing.....	56
4.3.6	Stopword Removal.....	56
4.3.7	Stemming.....	56
4.3.8	Normalisasi Slang.....	56
4.4	Balancing Data.....	57
4.5	Pelabelan Manual.....	57
4.6	Pembobotan TF-IDF.....	59
4.7	Pembagian Data (Train-Test Split).....	60
4.8	Pelatihan Model & Hyperparameter Tuning.....	60
4.8.1	Naive Bayes.....	60
4.8.2	Support Vector Machine (SVM).....	61
4.8.3	Penyimpanan Model Terbaik.....	62
4.9	Evaluasi Model.....	62
4.9.1	Confusion Matrix.....	62
4.9.2	Pengukuran Accuracy, Precision, Recall, dan F1-Score.....	66
4.10	Word Cloud.....	67
4.10.1	Word Cloud Semua Sentimen.....	67
	Perbandingan Kedua Word Cloud.....	68
4.10.2	Word Cloud Sentimen Positif.....	69
4.10.3	Word Cloud Sentimen Negatif.....	71
4.10.4	Word Cloud Sentimen Netral.....	73
4.11	Pie Chart.....	75
4.11.1	Pie Chart Sentimen Dataset Awal.....	75
4.11.2	Pie Chart Sentimen Dataset Updated.....	76
4.12	Hasil Klasifikasi Dan Perbandingan Hasil.....	77

4.12.1	Hasil Klasifikasi Dataset Awal (Naive Bayes & SVM).....	77
4.12.2	Tabel Akurasi Hasil (Naive Bayes & SVM).....	78
4.12.4	Akurasi Hasil (Naive Bayes & SVM).....	80
4.12.5	Keseluruhan Akurasi Hasil (Naive Bayes & SVM).....	81
4.13	Perbandingan Akurasi Naive Bayes dan SVM.....	82
4.1	Evolusi Opini Publik.....	83
BAB V PENUTUP		83
5.1	Kesimpulan.....	83
5.2	Saran.....	83
REFERENSI.....		85
LAMPIRAN.....		87



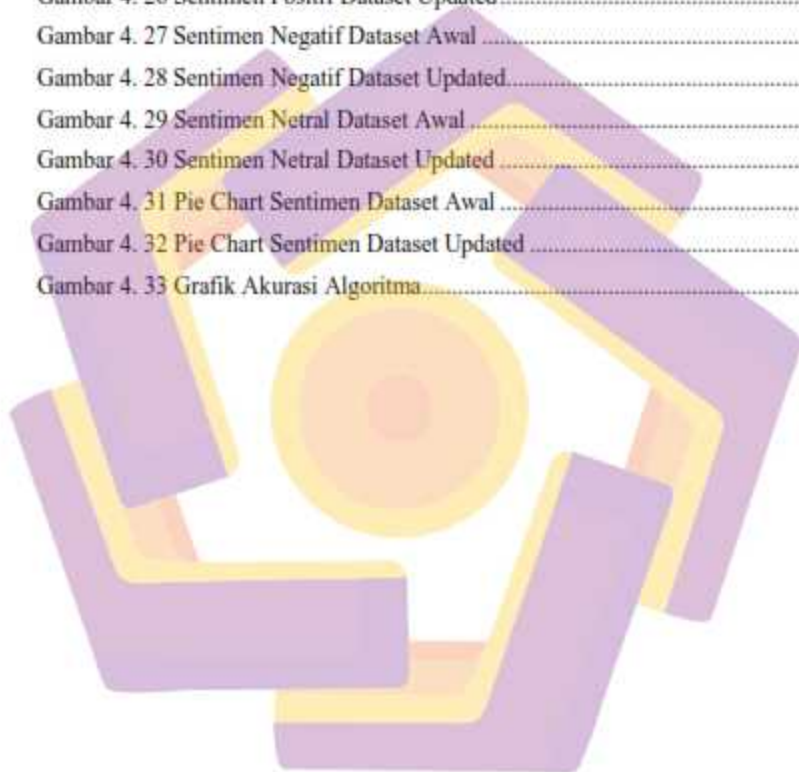
DAFTAR TABEL

Tabel 2 1 Keaslian Penelitian.....	15
Tabel 3 1 Statistik Dataset Updated.....	30
Tabel 3 2 Statistik Dataset Awal.....	31
Tabel 3 3 Hasil Cleaning Data.....	42
Tabel 3 4 Hasil Case Folding.....	44
Tabel 3 5 Hasil Tokenizing.....	45
Tabel 3 6 Hasil Stopword Removal.....	47
Tabel 4. 1 Akurasi Hasil Dataset Awal.....	76
Tabel 4. 2 Akurasi Hasil Dataset Updated.....	79
Tabel 4. 3 Keseluruhan Hasil Akhir.....	79
Tabel 4. 4 Evolusi Opini Publik.....	81

DAFTAR GAMBAR

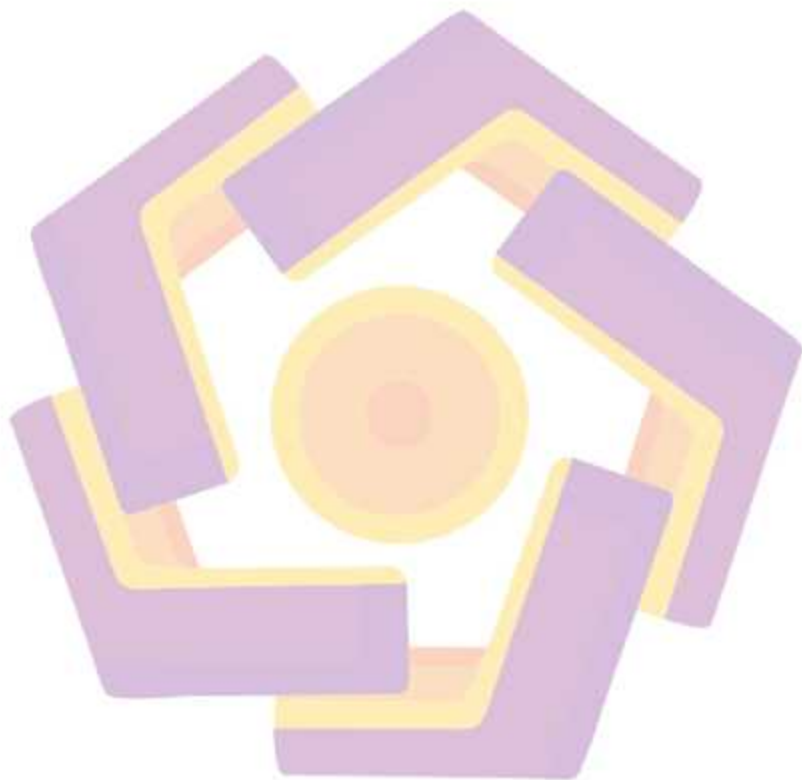
Gambar 3 1 Alur Penelitian.....	26
Gambar 3 2 Alur Scrapping Data.....	38
Gambar 3 3 Alur Filtering Bahasa	39
Gambar 3 4 Alur Cleaning Data.....	41
Gambar 3 5 Alur Case Folding	43
Gambar 3 6 Alur Tokenizing.....	44
Gambar 3 7 Alur Stopword Removal.....	46
Gambar 3 8 Alur Stemming	47
Gambar 4. 1 Source Code API YouTube.....	52
Gambar 4. 2 Save Data.....	53
Gambar 4. 3 Membaca Dataset.....	53
Gambar 4. 4 Source Code Filtering Bahasa	54
Gambar 4. 5 Source Code Cleaning Data.....	55
Gambar 4. 6 Source Code Case Folding	55
Gambar 4. 7 Source Code Tokenizing	56
Gambar 4. 8 Source Code Stopword Removal	56
Gambar 4. 9 Source Code Stemming	56
Gambar 4. 10 Source Code Normalisasi Slang.....	57
Gambar 4. 11 Source Code Balancing Data.....	57
Gambar 4. 12 Label Dataset Awal	58
Gambar 4. 13 Label Dataset Updated	58
Gambar 4. 14 Source Code TF IDF	59
Gambar 4. 15 Source Code Data Split	60
Gambar 4. 16 Source Code Training Model NB.....	61
Gambar 4. 17 Source Code Training Model SVM.....	61
Gambar 4. 18 Source Code Save Model	62
Gambar 4. 19 Hasil Confusion Matrix Naïve Bayes Dataset Awal	62
Gambar 4. 20 Hasil Confusion Matrix Naïve Bayes Dataset Updated	63

Gambar 4. 21 Hasil Confusion Matrix SVM Dataset Awal.....	63
Gambar 4. 22 Hasil Confusion Matrix SVM Dataset Updated.....	63
Gambar 4. 23 WordCloud Sentimen Dataset Awal	65
Gambar 4. 24 WordCloud Sentimen Dataset Updated	66
Gambar 4. 25 Sentimen Positif Dataset Awal.....	67
Gambar 4. 26 Sentimen Positif Dataset Updated.....	68
Gambar 4. 27 Sentimen Negatif Dataset Awal	69
Gambar 4. 28 Sentimen Negatif Dataset Updated.....	70
Gambar 4. 29 Sentimen Netral Dataset Awal	71
Gambar 4. 30 Sentimen Netral Dataset Updated	72
Gambar 4. 31 Pie Chart Sentimen Dataset Awal	73
Gambar 4. 32 Pie Chart Sentimen Dataset Updated	74
Gambar 4. 33 Grafik Akurasi Algoritma.....	80

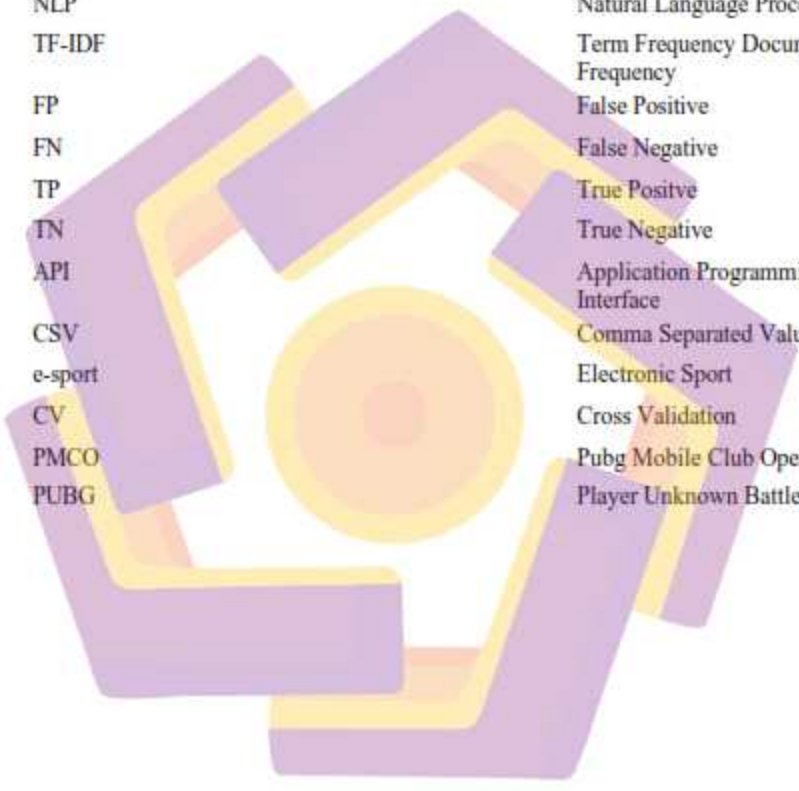


DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian	10
Lampiran 2. Dokumentasi Penelitian	11



DAFTAR LAMBANG DAN SINGKATAN



NBC	Naïve Bayes Clasifiers
SVM	Support Vector Machines
KNN	K-Nearest Neighbor
NLP	Natural Language Processing
TF-IDF	Term Frequency Document Frequency
FP	False Positive
FN	False Negative
TP	True Positive
TN	True Negative
API	Application Programming Interface
CSV	Comma Separated Values
e-sport	Electronic Sport
CV	Cross Validation
PMCO	Pubg Mobile Club Open
PUBG	Player Unknown Battle Ground

DAFTAR ISTILAH

Sentiment Analysis

Teknik untuk mengidentifikasi dan mengklasifikasikan opini atau emosi dalam teks ke dalam kategori positif, negatif, atau netral.

Text Mining

Proses penggalian informasi dan pola dari data teks yang tidak terstruktur.

Natural Language Processing (NLP)

Bidang kecerdasan buatan yang mempelajari interaksi antara komputer dan bahasa manusia.

Case Folding

Tahapan preprocessing untuk mengubah seluruh huruf pada teks menjadi huruf kecil.

Tokenizing

Proses memecah teks atau kalimat menjadi unit-unit kata (token).

Stopword Removal

Proses penghapusan kata-kata umum yang tidak memiliki pengaruh signifikan terhadap makna teks.

Stemming

Proses mengubah kata berimbuhan menjadi kata dasar.

TF-IDF

Metode pembobotan kata untuk menentukan tingkat kepentingan suatu kata dalam dokumen.



Term Frequency (TF)	Jumlah kemunculan suatu kata dalam satu dokumen.
Document Frequency (DF)	Jumlah dokumen yang mengandung suatu kata tertentu.
Naive Bayes Classifier (NBC)	Algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dengan asumsi independensi antar fitur.
Support Vector Machine (SVM)	Algoritma klasifikasi yang bekerja dengan mencari hyperplane optimal untuk memisahkan kelas data.
Confusion Matrix	Tabel evaluasi yang digunakan untuk mengukur kinerja model klasifikasi berdasarkan prediksi dan data aktual.
Accuracy	Ukuran kinerja model yang menunjukkan prediksi yang benar.
Precision	Ukuran ketepatan model dalam memprediksi suatu kelas tertentu.
Recall	Ukuran kemampuan model dalam menemukan kembali seluruh data yang relevan.

F1-Score	Rata-rata harmonis antara precision dan recall.
Word Cloud	Visualisasi data teks yang menampilkan kata berdasarkan frekuensi kemunculannya.
F1-Score	Rata-rata harmonis antara precision dan recall.
YouTube Data API	Antarmuka pemrograman aplikasi yang digunakan untuk mengakses data publik YouTube, termasuk komentar, metadata video, dan informasi pengguna.
Web Scraping	Teknik pengambilan data secara otomatis dari halaman web tanpa interaksi langsung dengan pengguna.
Train-Test Split	Teknik pembagian dataset menjadi data latih dan data uji untuk membangun dan mengevaluasi model klasifikasi.
Balancing Data	Proses penyeimbangan jumlah data pada setiap kelas sentimen agar model tidak bias terhadap kelas tertentu.
E-sport (Electronic Sport)	Bentuk kompetisi olahraga berbasis permainan digital yang dimainkan secara profesional.

INTISARI

YouTube merupakan salah satu platform media sosial yang paling banyak digunakan untuk berbagi konten digital, termasuk turnamen *e-sport* seperti PUBG Mobile PMCO Global 2019 Final. Keberhasilan tim Indonesia pada ajang tersebut menghasilkan berbagai opini dan reaksi dari penonton yang terekam melalui kolom komentar. Penelitian ini bertujuan menganalisis sentimen komentar pengguna YouTube terkait turnamen tersebut dengan menggunakan metode Naive Bayes Classifier (NBC) sebagai algoritma utama dan Support Vector Machine (SVM) sebagai model pembandingan.

Data penelitian diperoleh melalui web scraping terhadap seluruh komentar yang dipublikasikan sejak video diunggah hingga tahun 2025. Tahapan pengolahan data meliputi filtering bahasa, pembersihan komentar, serta preprocessing teks yang mencakup case folding, tokenizing, stopword removal, stemming, dan normalisasi slang. Selanjutnya, pembobotan fitur dilakukan menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), sebelum data digunakan untuk pelatihan dan pengujian model dengan beberapa skenario pembagian data.

Evaluasi performa model dilakukan menggunakan confusion matrix serta metrik accuracy, precision, recall, dan F1-score. Penelitian ini diharapkan memberikan gambaran komprehensif mengenai pola sentimen publik terhadap turnamen *e-sport* serta menjadi referensi dalam pengembangan analisis sentimen menggunakan metode pembelajaran mesin pada data komentar YouTube.

Kata kunci: Analisis Sentimen, YouTube, PUBG Mobile, Naive Bayes, SVM, TF-IDF.

ABSTRACT

YouTube is one of the most widely used social media platforms for sharing digital content, including e-sports tournaments such as the PUBG Mobile PMCO Global 2019 Final. The success of the Indonesian team at the event generated various opinions and reactions from viewers, which were recorded in the comments section. This study aims to analyze the sentiment of YouTube users' comments related to the tournament using the Naïve Bayes Classifier (NBC) method as the main algorithm and Support Vector Machine (SVM) as a comparison model.

The research data was obtained through web scraping of all comments published since the video was uploaded until 2025. Data processing stages included language filtering, comment cleaning, and text preprocessing, which encompassed case folding, tokenizing, stopword removal, stemming, and slang normalization. Next, feature weighting was performed using Term Frequency-Inverse Document Frequency (TF-IDF), before the data was used for model training and testing with several data partitioning scenarios.

Model performance evaluation was conducted using a confusion matrix and accuracy, precision, recall, and F1-score metrics. This study is expected to provide a comprehensive overview of public sentiment patterns toward e-sports tournaments and serve as a reference in the development of sentiment analysis using machine learning methods on YouTube comment data.

Keyword: Sentiment Analysis, YouTube, PUBG Mobile, Naive Bayes, SVM, TF-IDF.