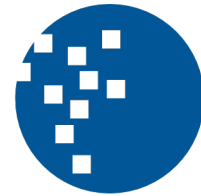


**ANALISIS SENTIMEN TERHADAP PELATIH BARU
TIMNAS INDONESIA DI INSTAGRAM MENGGUNAKAN
ALGORITMA NAIVE BAYES**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

MARSELINO LENGYU PANTOUW
00000061909

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

**ANALISIS SENTIMEN TERHADAP PELATIH BARU
TIMNAS INDONESIA DI INSTAGRAM MENGGUNAKAN
ALGORITMA NAIVE BAYES**



SKRIPSI

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

**MARSELINO LENGYU PANTOUW
00000061909**

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Marselino Lengyu Pantouw
Nomor Induk Mahasiswa : 00000061909
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen terhadap Pelatih Baru Timnas Indonesia di Instagram Menggunakan Algoritma 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, 3 Juli 2025



(Marselino Lengyu Pantouw)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN TERHADAP PELATIH BARU TIMNAS INDONESIA DI INSTAGRAM MENGGUNAKAN ALGORITMA NAIVE BAYES

oleh

Nama : Marselino Lengyu Pantouw
NIM : 00000061909
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Jumat, 18 Juli 2025

Pukul 08.00 s/s 10.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Wirawan Istiono, S.Kom., M.Kom.)

NIDN: 0313048304

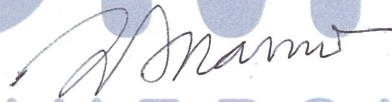
Penguji



(Sy. Yuliani Yakub, S.Kom., M.T., Ph.D.)

NIDN: 0411037904

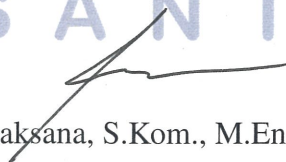
Pembimbing



(Dr. Ir. P. M. Winarno, M.Kom.)

NIDN: 0330106002

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 : Marselino Lengyu Pantouw
NIM : 00000061909
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen terhadap Pelatih
Baru Timnas Indonesia di Instagram
Menggunakan Algoritma 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, 3 Juli 2025

Yang menyatakan

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



Marselino Lengyu Pantouw

**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

"Stay low, stay quiet, keep it simple, don't expect too much, enjoy what you have."

- Dean Koontz

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan tugas akhir ini dengan judul: Analisis Sentimen terhadap Pelatih Baru Timnas Indonesia di Instagram Menggunakan Algoritma Naive Bayes dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir 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. Bapak Dr. Ir. P. M. Winarno, M.Kom., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Pacar saya yang telah membantu dan mendukung dalam berbagai bentuk, sehingga penulis dapat terus berjuang menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 Juli 2025



Marselino Lengyu Pantouw

**ANALISIS SENTIMEN TERHADAP PELATIH BARU TIMNAS
INDONESIA DI INSTAGRAM MENGGUNAKAN ALGORITMA NAIVE
BAYES**

Marselino Lengyu Pantouw

ABSTRAK

Perkembangan media sosial telah menjadikan platform seperti Instagram sebagai wadah masyarakat untuk menyampaikan opini, termasuk terhadap isu olahraga nasional. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap pelatih baru Tim Nasional Sepak Bola Indonesia pasca pergantian dari Shin Tae-yong ke Patrick Kluivert. Dataset diambil dari komentar pada pertandingan Indonesia melawan Australia dan Bahrain. Proses pengolahan data meliputi *preprocessing*, pelabelan menggunakan *TextBlob*, serta klasifikasi menggunakan algoritma *Multinomial Naïve Bayes* dengan metode ekstraksi fitur TF-IDF dan penyeimbangan data menggunakan *SMOTE*. Evaluasi dilakukan pada tiga skenario pembagian data: rasio 70:30 dengan akurasi 85,77%, rasio 80:20 dengan akurasi 85,88%, dan rasio 90:10 dengan akurasi tertinggi 88,39%. Hasil ini membuktikan bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan sentimen secara konsisten dan efektif dalam konteks komentar media sosial terkait isu olahraga.

Kata kunci: Analisis Sentimen, Instagram, *Multinomial Naïve Bayes*, *Natural Language Processing*, *TextBlob*.



***SENTIMENT ANALYSIS TOWARDS THE NEW COACH OF INDONESIA
NATIONAL FOOTBALL TEAM ON INSTAGRAM USING NAIVE BAYES
ALGORITHM***

Marselino Lengyu Pantouw

ABSTRACT

The growth of social media has positioned platforms like Instagram as a space for the public to express opinions, including those related to national sports issues. This study aims to analyze public sentiment toward the new coach of the Indonesian National Football Team following the transition from Shin Tae-yong to Patrick Kluivert. The dataset was collected from Instagram comments on Indonesia's matches against Australia and Bahrain. The data processing involved preprocessing, sentiment labeling using TextBlob, and classification using the Multinomial Naïve Bayes algorithm with TF-IDF for feature extraction and SMOTE for class balancing. Evaluation was conducted across three data-splitting scenarios: 70:30 with an accuracy of 85.77%, 80:20 with 85.88%, and 90:10 with the highest accuracy of 88.39%. These results demonstrate that the Naïve Bayes algorithm consistently and effectively classifies sentiment in social media comments related to sports issues.

Keywords: Instagram, Multinomial Naïve Bayes, Natural Language Processing, Sentiment Analysis, TextBlob.



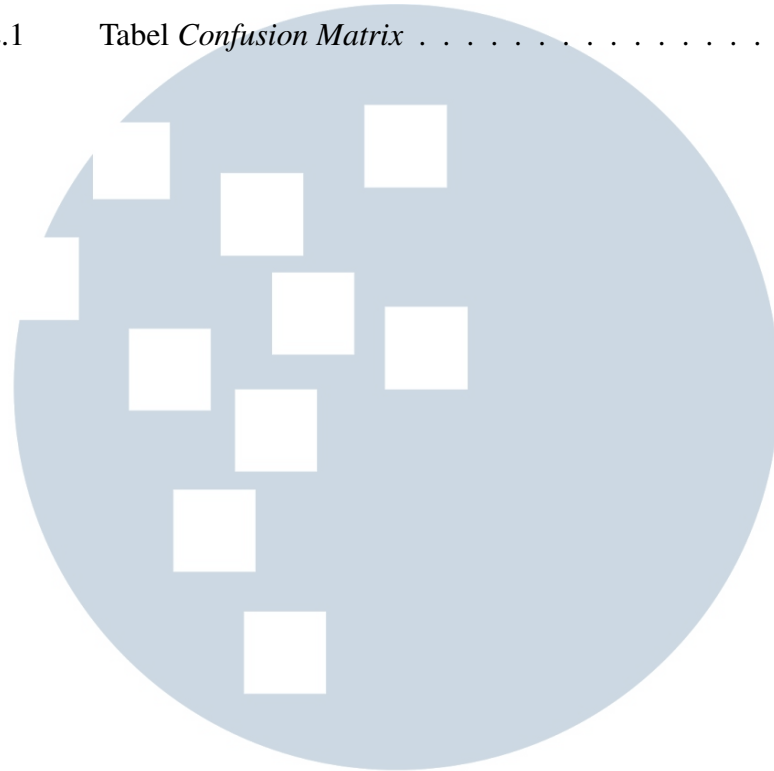
DAFTAR ISI

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	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.5.1 Manfaat Teoritis	4
1.5.2 Manfaat Praktis	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Analisis Sentimen	6
2.2 Instagram	6
2.3 Preprocessing	6
2.3.1 Cleaning Data	7
2.3.2 Case Folding	7
2.3.3 Tokenization	7
2.3.4 Normalization	7
2.3.5 Stopword Removal	8
2.3.6 Stemming	8
2.4 TF-IDF (<i>Term Frequency – Inverse Document Frequency</i>)	8
2.4.1 Term Frequency	8
2.4.2 Inverse Document Frequency	9
2.4.3 TF-IDF	9
2.5 SMOTE	10
2.6 Naïve Bayes	10
2.6.1 Multinomial Naïve Bayes	11
2.6.2 Langkah Umum Klasifikasi Menggunakan <i>Naïve Bayes</i>	11
2.6.3 Keunggulan Naïve Bayes	12
2.6.4 Kelemahan Naïve Bayes	12
2.7 Confusion Matrix	12
2.7.1 Accuracy	13
2.7.2 Precision	13
2.7.3 Recall	13
2.7.4 F1-Score	14

BAB 3	METODE PENELITIAN	15
3.1	Metode Penelitian	15
3.2	Alur Penelitian	15
3.3	Pengumpulan Data	16
3.4	Spesifikasi Sistem	17
3.4.1	Perangkat Keras	18
3.4.2	Perangkat Lunak	18
3.5	Perancangan Sistem	18
3.5.1	Preprocessing Data	18
3.5.2	Labeling Data	20
3.5.3	TF-IDF Vectorization	21
3.5.4	SMOTE (Oversampling Data)	21
3.5.5	Splitting Data (Train-Test Split)	22
3.5.6	Pelatihan Model Multinomial Naive Bayes	22
3.6	Evaluasi Model	22
BAB 4	ANALISIS DAN HASIL PENELITIAN	23
4.1	Implementasi Sistem	23
4.2	Pengumpulan Data	23
4.3	Preprocessing Data	24
4.3.1	Cleaning Data	24
4.3.2	Case Folding	25
4.3.3	Tokenization	26
4.3.4	Normalization	27
4.3.5	Stopword Removal	28
4.3.6	Stemming	29
4.3.7	Finalisasi Teks	29
4.4	Labeling Data	30
4.5	TF-IDF Vectorization	31
4.6	SMOTE	32
4.7	Splitting Data	33
4.8	Pelatihan Model Multinomial Naive Bayes	35
4.9	Evaluasi Model	36
4.9.1	Confusion Matrix	37
4.9.2	Classification Report	40
4.10	Visualisasi Sentimen	42
4.10.1	Distribusi Sentimen (Bar Chart)	43
4.10.2	WordCloud	43
BAB 5	SIMPULAN DAN SARAN	46
5.1	Simpulan	46
5.2	Saran	46
DAFTAR PUSTAKA		48

DAFTAR TABEL

Tabel 2.1	Tabel <i>Confusion Matrix</i>	12
-----------	---	----



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 3.1	Flow Alur Penelitian	16
Gambar 3.2	Alur Pengumpulan Komentar di Instagram	17
Gambar 3.3	Flow <i>Preprocessing</i> Data	19
Gambar 3.4	Flow Labeling Data menggunakan TextBlob	21
Gambar 4.1	Hasil Importing Dataset Sebelum di Proses	24
Gambar 4.2	Hasil Cleaning Data Komentar Instagram	25
Gambar 4.3	Hasil Case Folding Komentar Instagram	26
Gambar 4.4	Hasil Tokenizing Komentar Instagram	26
Gambar 4.5	Hasil Stopword Removal pada Komentar Instagram	28
Gambar 4.6	Hasil Akhir Teks setelah Preprocessing	30
Gambar 4.7	Hasil Pelabelan Komentar Instagram	31
Gambar 4.8	Distribusi Label Sentimen (Netral, Positif, Negatif)	31
Gambar 4.9	Dimensi Vektor Hasil TF-IDF	32
Gambar 4.10	Distribusi Label setelah Penerapan SMOTE	33
Gambar 4.11	Jumlah Data Latih dan Uji setelah Pembagian	35
Gambar 4.12	Hasil Visualisasi Confusion Matrix Rasio 70:30	38
Gambar 4.13	Hasil Visualisasi Confusion Matrix Rasio 80:20	39
Gambar 4.14	Hasil Visualisasi Confusion Matrix Rasio 90:10	39
Gambar 4.15	Hasil Visualisasi Classification Report Rasio 70:30	41
Gambar 4.16	Hasil Visualisasi Classification Report Rasio 80:20	41
Gambar 4.17	Hasil Visualisasi Classification Report Rasio 90:10	42
Gambar 4.18	Distribusi Komentar berdasarkan Sentimen	43
Gambar 4.19	Hasil WordCloud untuk Sentimen Positif	44
Gambar 4.20	Hasil WordCloud untuk Sentimen Negatif	44
Gambar 4.21	Hasil WordCloud untuk Sentimen Netral	45



DAFTAR KODE

Kode 3.1	Fungsi pelabelan berdasarkan polaritas TextBlob	20
Kode 4.1	Importing Dataset ke dalam Jupyter Notebook	23
Kode 4.2	Proses Cleaning Data Komentar	24
Kode 4.3	Proses Case Folding Komentar	25
Kode 4.4	Proses Tokenizing Komentar	26
Kode 4.5	Proses Normalisasi Menggunakan Kamus Sederhana	27
Kode 4.6	Proses Stopword Removal menggunakan NLTK	28
Kode 4.7	Proses Stemming dengan Library Sastrawi	29
Kode 4.8	Penggabungan Kata-Kata menjadi Kalimat	29
Kode 4.9	Proses Penentuan Label Sentimen dengan TextBlob	30
Kode 4.10	Proses Ekstraksi Fitur Menggunakan TF-IDF	32
Kode 4.11	Implementasi SMOTE pada Data TF-IDF	33
Kode 4.12	Pembagian Data Latih dan Uji menggunakan <code>train_test_split</code> . . .	34
Kode 4.13	Pelatihan Model dan Prediksi menggunakan <i>Multinomial Naïve Bayes</i>	36
Kode 4.14	Visualisasi Confusion Matrix	37
Kode 4.15	Pembuatan Classification Report	40



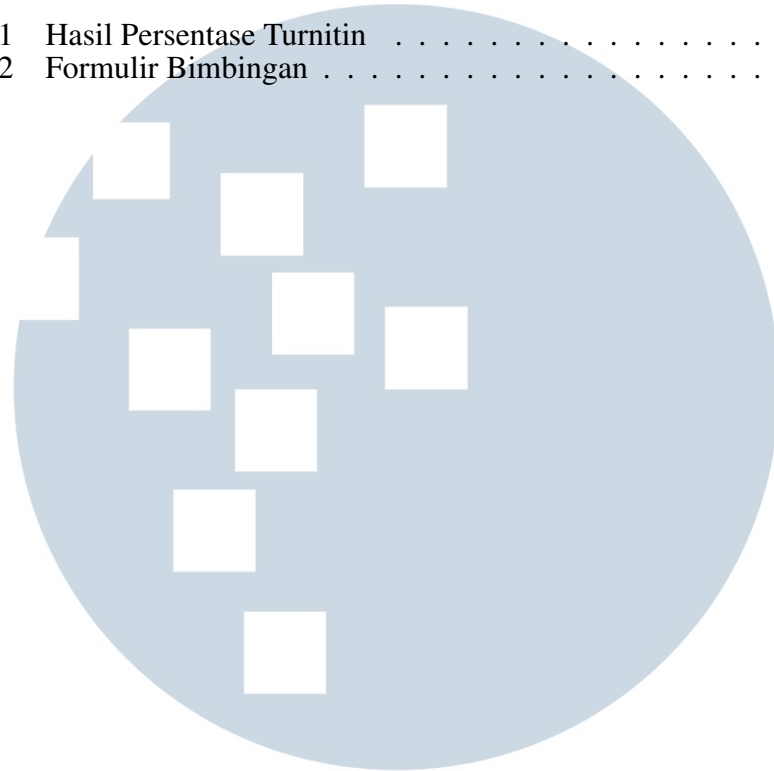
DAFTAR RUMUS

Rumus 2.1	<i>Term Frequency</i>	9
Rumus 2.2	<i>Inverse Document Frequency</i>	9
Rumus 2.3	<i>TF-IDF</i>	9
Rumus 2.4	<i>Multinomial Naïve Bayes</i>	11
Rumus 2.5	<i>Accuracy</i>	13
Rumus 2.6	<i>Precision</i>	13
Rumus 2.7	<i>Recall</i>	14
Rumus 2.8	<i>F1-Score</i>	14



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	51
Lampiran 2	Formulir Bimbingan	53



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA