

**PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES
DENGAN ADABOOST UNTUK ANALISIS OPINI PUBLIK
TENTANG NATURALISASI PEMAIN SEPAK BOLA DI
INDONESIA**

SKRIPSI

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



disusun oleh

MOCHAMMAD BIMA HARTONO

19.12.1387

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES
DENGAN ADABOOST UNTUK ANALISIS OPINI PUBLIK
TENTANG NATURALISASI PEMAIN SEPAK BOLA DI
INDONESIA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Sistem Informasi



disusun oleh

MOCHAMMAD BIMA HARTONO

19.12.1387

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES DENGAN
ADABOOST UNTUK ANALISIS OPINI PUBLIK TENTANG NATURALISASI
PEMAIN SEPAK BOLA DI INDONESIA

yang disusun dan diajukan oleh

MOCHAMMAD BIMA HARTONO

19.12.1387

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Agustus 2025

Dosen Pembimbing,



Sharazita Dyah Anggita, M.Kom
NIK : 190302285

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES DENGAN
ADABOOST UNTUK ANALISIS OPINI PUBLIK TENTANG NATURALISASI
PEMAIN SEPAK BOLA DI INDONESIA

yang disusun dan diajukan oleh

MOCHAMMAD BIMA HARTONO

19.12.1387

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Yoga Pristvanto, S.Kom., M.Eng.
NIK. 190302412

Irma Rofni Wulandari, S.Pd., M.Eng.
NIK. 190302329

Sharazita Dyah Anggita, S.Kom., M.Kom.
NIK. 190302285

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : MOCHAMMAD BIMA HARTONO
NIM : 19.12.1387

Menyatakan bahwa Skripsi dengan judul berikut:

**PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES
DENGAN ADABOOST UNTUK ANALISIS OPINI PUBLIK TENTANG
NATURALISASI PEMAIN SEPAK BOLA DI INDONESIA**

Dosen Pembimbing : Sharazita Dyah Anggita, 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, 25 Agustus 2025



Mochammad Bima Hartono

HALAMAN PERSEMBAHAN

Segala puji serta syukur saya panjatkan ke hadirat Allah SWT, atas limpahan rahmat, kasih sayang, dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi ini yang berjudul "**PERBANDINGAN ALGORITMA MULTINOMIAL NAIVE BAYES DENGAN ADABOOST UNTUK ANALISIS OPINI PUBLIK TENTANG NATURALISASI PEMAIN SEPAK BOLA DI INDONESIA**".

Karya ini adalah bukti dari usaha dan perjalanan yang penuh dengan tantangan selama masa perkuliahan.

Dengan segala kerendahan hati, karya ini saya persembahkan kepada:

1. Saya sendiri, yang telah berjuang melewati berbagai rintangan selama masa studi. Saya bangga atas ketekunan dan semangat yang tidak pernah padam.
2. Kedua orang tua tercinta, yang selalu menjadi sumber inspirasi, kasih sayang, doa, dan pengorbanan tanpa batas.
3. Keluarga dan saudara-saudara, yang senantiasa memberikan dukungan moral dan motivasi dalam setiap langkah perjalanan saya.
4. Ibu Sharazita Dyah Anggita, M.Kom selaku dosen pembimbing, yang telah sabar memberikan bimbingan, arahan, serta ilmu yang berharga selama proses penyusunan skripsi ini.
5. Teman-teman dan sahabat seperjuangan, yang telah menemani, memberi semangat, dan menjadi bagian dari perjalanan hidup saya selama masa perkuliahan hingga penyelesaian skripsi.
6. Almamater tercinta, Universitas AMIKOM Yogyakarta, tempat saya menimba ilmu dan berkembang. Semoga karya ini dapat memberikan kontribusi bagi kemajuan ilmu pengetahuan di masa depan.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul *“Perbandingan Algoritma Multinomial Naive Bayes dengan AdaBoost untuk Analisis Opini Publik tentang Naturalisasi Pemain Sepak Bola di Indonesia.”*

Tugas akhir ini disusun sebagai syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penulis menyampaikan terima kasih kepada:

1. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Prof. Dr. Kusri, M.Kom., selaku Dekan Fakultas Ilmu Komputer.
3. Anggit Dwi Hartanto, M.Kom., selaku Kaprodi Sistem Informasi.
4. Sharazita Dyah Anggita, M.Kom., selaku Dosen Pembimbing.
5. Seluruh dosen dan staf Universitas Amikom Yogyakarta.
6. Kedua orang tua dan keluarga atas dukungan dan doa.
7. Kepada istri dan anak terkasih atas memotivasi dan dukungannya
8. Teman-teman seperjuangan serta seluruh pihak yang turut membantu dalam proses penyusunan tugas akhir ini.

Penulis menyadari tugas akhir ini masih memiliki kekurangan, sehingga kritik dan saran yang membangun sangat diharapkan. Semoga karya ini bermanfaat bagi pembaca serta dapat menjadi referensi dalam pengembangan ilmu, khususnya di bidang teknologi informasi dan analisis data.

Yogyakarta, 25 Agustus 2025

Mochammad Bima Hartono

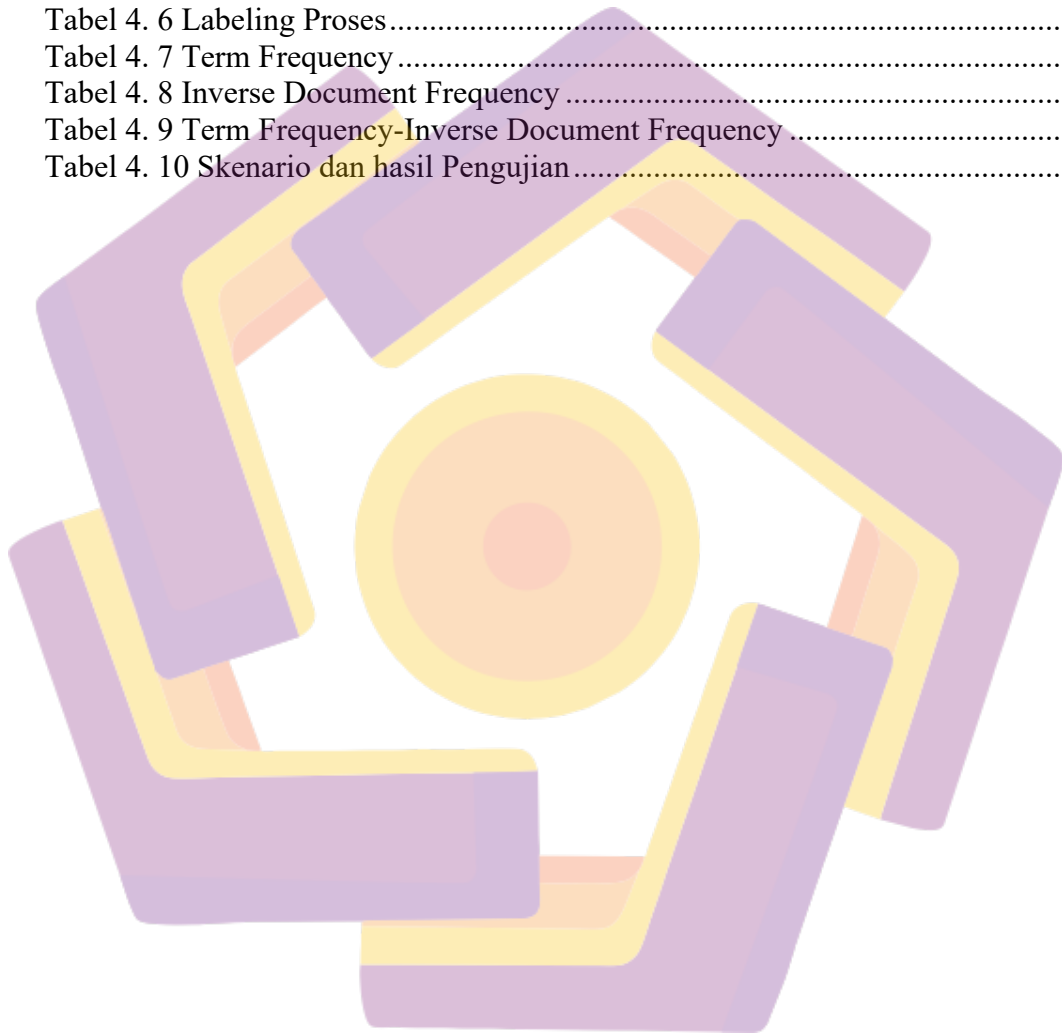
DAFTAR ISI

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.....	ix
DAFTAR GAMBAR.....	x
INTISARI	xi
<i>ABSTRACT</i>	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	9
2.1 Studi Literatur	9
2.2 Dasar Teori	18
2.2.1 Crawling Data Twitter	18
2.2.2 Preprocessing	18
2.2.3 Labelling	21
2.2.4 Ekstraksi Fitur dan Pembobotan Kata.....	22
2.2.5 Multinomial Naïve Bayes	23
2.2.6 Adaboost	25
2.2.7 Evaluasi dan validasi	26
BAB III METODE PENELITIAN	29
3.1 Objek Penelitian.....	29

3.2 Alur Penelitian	30
3.3 Analisis kebutuhan perangkat	31
3.4 Uraian Metodologi	31
BAB IV HASIL DAN PEMBAHASAN	34
4.1 Crawling	34
4.2 Preprocessing	35
4.2.1 Case Folding	35
4.2.2 Punctuation	36
4.2.3 Stopwords Removal	38
4.2.4 Tokenization	40
4.2.5 Stemming	41
4.2.6 Labelling	42
4.2.7 TF-IDF	43
4.3 Hasil Confusion matrix	47
4.4 Pembahasan	51
BAB V PENUTUP	54
5.1 Kesimpulan	54
5.2 Saran	54
Daftar Pustaka	56

DAFTAR TABEL

Tabel 4. 1 Case folding	36
Tabel 4. 2 Clean Punctuation	37
Tabel 4. 3 Stopwords Removals	38
Tabel 4. 4 Proses Tokenizing	40
Tabel 4. 5 Proses Stemming	41
Tabel 4. 6 Labeling Proses	42
Tabel 4. 7 Term Frequency	43
Tabel 4. 8 Inverse Document Frequency	44
Tabel 4. 9 Term Frequency-Inverse Document Frequency	46
Tabel 4. 10 Skenario dan hasil Pengujian	49



DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	30
Gambar 3. 2 Cross Validation.....	33
Gambar 4. 1 Autentikasi Token Twitter	34
Gambar 4. 2 Hasil Crawling Data Keseluruhan.....	35
Gambar 4. 3 Kode case folding.....	36
Gambar 4. 4 Kode Program Clean Punctuation.....	38
Gambar 4. 5 Kode Program Stopword Removal	39
Gambar 4. 6 Kode Program Tokenizing	41
Gambar 4. 7 Confusion Matrix Hasil Skenario Data 90:10	47
Gambar 4. 8 Confusion Matrix Hasil Skenario Data 80:20	48
Gambar 4. 9 Confusion Matrix Hasil Skenario Data 70:30	49
Gambar 4. 10 Hasil Cross Validation	50

INTISARI

Penelitian ini bertujuan untuk menganalisis sentimen publik pada media sosial Twitter mengenai isu naturalisasi pemain sepak bola di Indonesia dengan membandingkan kinerja dua algoritma klasifikasi: *Multinomial Naïve Bayes* (MNB) dan AdaBoost. Data diperoleh melalui *crawling* tweet sepanjang tahun 2024 dengan kata kunci "sepakbola", "timnasional", dan "naturalisasi". Data mentah kemudian melalui serangkaian tahap *preprocessing* yang meliputi *case folding*, penghapusan tanda baca, *tokenizing*, *stopword removal*, dan *stemming*. Pelabelan sentimen (positif, negatif, netral) dilakukan secara otomatis menggunakan model transformer, dan fitur diekstraksi menggunakan TF-IDF. Hasil evaluasi menunjukkan bahwa model *Multinomial Naïve Bayes* secara konsisten menunjukkan performa yang lebih unggul, berhasil mencapai akurasi hingga 82%. Sebaliknya, penerapan AdaBoost menghasilkan akurasi yang lebih rendah secara signifikan, yaitu hanya 65.29% dalam pengujian *cross-validation*. Kedua model mengalami kesulitan dalam mengklasifikasikan sentimen netral secara akurat karena ketidakseimbangan data. Penelitian ini menyimpulkan bahwa untuk klasifikasi teks dari Twitter yang cenderung memiliki noise tinggi, model MNB yang lebih sederhana terbukti lebih efektif dibandingkan AdaBoost.

Kata kunci: Analisis Sentimen, Naïve Bayes, AdaBoost, Naturalisasi Pemain, Twitter

ABSTRACT

This study aims to analyze public sentiment on Twitter regarding the issue of football player naturalization in Indonesia by comparing the performance of two classification algorithms: Multinomial Naïve Bayes (MNB) and AdaBoost. Data was obtained by crawling tweets throughout 2024 with the keywords "football", "national team", and "naturalization". The raw data underwent a preprocessing stage, including case folding, punctuation removal, tokenizing, stopword removal, and stemming. Sentiment labeling (positive, negative, and neutral) was performed automatically using a transformer model, and features were extracted using TF-IDF. The evaluation results show that the Multinomial Naïve Bayes model consistently demonstrated superior performance, achieving a classification accuracy of up to 82%. In contrast, the implementation of AdaBoost yielded a significantly lower accuracy of 65.29% in cross-validation testing. Both models struggled to accurately classify neutral sentiment due to a data imbalance. This study concludes that for text classification on noisy Twitter data, the simpler MNB model proved to be more effective than AdaBoost.

Keywords: Sentiment Analysis, Naïve Bayes, AdaBoost, Player Naturalization, Twitter.