

**KOMPARASI ALGORITMA *NAIVE BAYES* DAN *LOGISTIC REGRESSION* TERHADAP ANALISIS SENTIMEN
PERGANTIAN PELATIH TIMNAS INDONESIA DI MEDIA
SOSIAL X (*TWITTER*)**

SKRIPSI

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:
BUSTAMI NUR QALBI
NIM D1041181013



FAKULTAS TEKNIK
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Bustami Nur Qalbi

NIM : D1041181013

menyatakan bahwa dalam skripsi yang berjudul “Komparasi Algoritma *Naive Bayes* dan *Logistic Regression* Terhadap Analisis Sentimen Pergantian Pelatih Timnas Indonesia di Media Sosial X (*Twitter*)” tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi manapun. Sepanjang pengetahuan Saya, tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Demikian pernyataan ini dibuat dengan sebenar-benarnya. Saya sanggup menerima konsekuensi akademis dan hukum di kemudian hari apabila pernyataan yang dibuat ini tidak benar.

Pontianak, 25 Juni 2025

Bustami Nur Qalbi
NIM D1041181013



KEMENTERIAN PENDIDIKAN TINGGI, SAINS,
DAN TEKNOLOGI
UNIVERSITAS TANJUNGPURA
FAKULTAS TEKNIK

Jalan Prof. Dr. H. Hadari Nawawi Pontianak 78124
Telepon (0561) 740186, WA: +6282152280907
Email : ft@untan.ac.id Website : <http://teknik.untan.ac.id>

HALAMAN PENGESAHAN

KOMPARASI ALGORITMA *NAIVE BAYES* DAN *LOGISTIC REGRESSION*
TERHADAP ANALISIS SENTIMEN PERGANTIAN PELATIH TIMNAS
INDONESIA DI MEDIA SOSIAL X (*TWITTER*)

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:

Bustami Nur Qalbi
NIM D1041181013

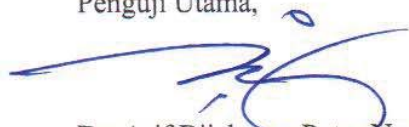
Telah dipertahankan di depan Penguji Skripsi pada tanggal 25 Juni 2025 dan
diterima sebagai salah satu persyaratan untuk memperoleh gelar sarjana.

Susunan Penguji Skripsi:

Ketua,


Prof. Dr. Herry Sujaini, S.T., M.T.
NIP 196806291997021001

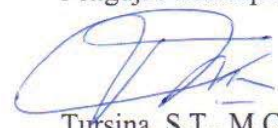
Penguji Utama,


Dr. Arif Bijaksana Putra Negara, S.T., M.T.
NIP 197208081998021002

Sekretaris,


Rina Septiriana, S.T., M.Cs.
NIP 198709232020122001

Penguji Pendamping,

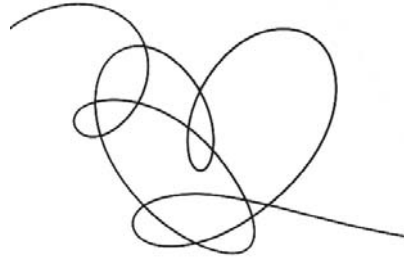

Tursina, S.T., M.Cs.
NIP 197801152002122003

Pontianak, 25 Juni 2025
Dekan,


Dr. Ing. Ir. Slamet Widodo, M.T., IPM
NIP 196712231992031002

HALAMAN PERSEMBAHAN

ㄥ ㄩ 結



Skripsi ini tidak didedikasikan kepada siapapun.

Di halaman ini saya ingin menuliskan ucapan terima kasih yang sebesar-besarnya kepada beberapa sosok.

Terima kasih untuk kedua orang tua saya yang terkhusus telah membiayai kuliah saya dari awal masuk kuliah sampai saya lulus.

Terima kasih kepada Ibu Rustinah yang telah memberikan tumpangan tempat tinggal dan juga mengurus kehidupan saya semasa kuliah.

Jeon Jung-kook dari BTS dengan lagu yang dinyanyikan olehnya berjudul 'Euphoria' yang telah menemani hari-hari saya semasa kuliah.

migi

Sepeda motor kesayangan, Honda Vario 150 eSP (2016) yang telah menemani perjalanan dan aktivitas saya semasa kuliah.

Ed Sheeran, aespa, BTS, YOASOBI, Twice, eclat story.

KATA PENGANTAR

Puji syukur penulis panjatkan Kepada Tuhan Yang Maha Esa, karena berkat dan rahmat-Nya penulis dapat menyelesaikan penelitian dengan judul “Komparasi Algoritma *Naive Bayes* dan *Logistic Regression* Terhadap Analisis Sentimen Pergantian Pelatih Timnas Indonesia di Media Sosial X (*Twitter*)”. Penelitian ini dibuat untuk memenuhi syarat dalam menyelesaikan kegiatan program studi sarjana Informatika Universitas Tanjungpura.

Penelitian ini menjelaskan tentang perbandingan performa algoritma *Naive Bayes* dan *Logistic Regression* terhadap analisis sentimen pergantian pelatih Timnas Indonesia di media sosial X. Penelitian ini bertujuan untuk membandingkan performa kedua algoritma berdasarkan matriks seperti akurasi, *precision*, *recall*, dan *F1-score*, serta memahami sentimen publik terhadap keputusan tersebut berdasarkan data *tweet* yang telah dikumpulkan.

Secara khusus penulis menyampaikan terima kasih yang sebesar-besarnya kepada Bapak Prof. Dr. Herry Sujaini, S.T., M.T., dan Ibu Rina Septiriana, S.T., M.T., selaku dosen pembimbing utama dan pembimbing pendamping tugas akhir saya, Bapak Dr. Arif Bijaksana Putra Negara, S.T., M.T., dan Ibu Tursina, S.T., M.Cs., selaku dosen penguji utama dan penguji pendamping tugas akhir saya, dan Ibu Eva Faja Ripanti, S.Kom., MMSI., Ph.D., selaku dosen pembimbing akademik saya, serta seluruh rekan-rekan mahasiswa di program studi Informatika Universitas Tanjungpura yang telah memberikan banyak *insight* dalam menyusun penelitian ini.

Penulis berharap penelitian ini dapat memberikan manfaat bagi masyarakat luas, khususnya dalam memahami persepsi publik terhadap isu tertentu melalui analisis sentimen di media sosial. Penulis menerima masukan yang konstruktif agar dapat menyempurnakan penelitian ini.

Pontianak, 25 Juni 2025
Penulis,

Bustami Nur Qalbi

ABSTRAK

Perubahan pelatih Timnas Indonesia pada awal 2025, khususnya penunjukan Patrick Kluivert, memicu beragam respons dari masyarakat yang disuarakan secara masif melalui media sosial X (sebelumnya *Twitter*). Penelitian ini bertujuan untuk membandingkan performa algoritma *Naive Bayes* dan *Logistic Regression* dalam menganalisis sentimen publik terhadap pergantian pelatih Timnas Indonesia, Patrick Kluivert, pada media sosial X selama periode 6 Januari hingga 6 Maret 2025. Data sebanyak 4.847 *tweet* berbahasa Indonesia dikumpulkan melalui teknik *crawling* menggunakan *Google Colab* dan diberi label sentimen (positif, negatif, netral) dengan bantuan empat platform AI: *ChatGPT*, *Copilot*, *Grok*, dan *Blackbox AI*. Proses *text-preprocessing* meliputi *case folding*, *cleaning*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*, diikuti ekstraksi fitur menggunakan TF-IDF. Untuk mengatasi ketidakseimbangan kelas, teknik SMOTE diterapkan. Pengujian dilakukan dengan delapan skenario berdasarkan variasi rasio data (70:30 dan 80:20) dan penggunaan SMOTE, dengan evaluasi performa menggunakan matriks akurasi, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa *Logistic Regression* dengan SMOTE menghasilkan performa terbaik dengan akurasi 83,80%, *precision* 82,80%, *recall* 82,70%, dan *F1-score* 82,80%. Analisis sentimen mengindikasikan dominasi sentimen negatif, mencerminkan kritik masyarakat terhadap pergantian pelatih. Penelitian ini memberikan wawasan tentang persepsi publik dan menunjukkan bahwa *Logistic Regression* lebih unggul dibandingkan *Naive Bayes* dalam konteks analisis sentimen berbasis data media sosial.

Kata kunci: Analisis Sentimen, Pergantian Pelatih, Patrick Kluivert, *Twitter/X*, *Naive Bayes*, *Logistic Regression*, SMOTE

ABSTRACT

The appointment of Patrick Kluivert as the new head coach of the Indonesian National Football Team in early 2025 sparked diverse public reactions, widely expressed on the social media platform X (formerly Twitter). This study aims to compare the performance of Naive Bayes and Logistic Regression algorithms in analyzing public sentiment toward the coaching change on X from January 6 to March 6, 2025. A dataset of 4,847 Indonesian-language tweets was collected through crawling using Google Colab and labeled for sentiment (positive, negative, neutral) with the assistance of four AI platforms: ChatGPT, Copilot, Grok, and Blackbox AI. The text preprocessing steps included case folding, cleaning, normalization, tokenizing, stopword removal, and stemming, followed by feature extraction using TF-IDF. To address class imbalance, the SMOTE technique was applied. Eight testing scenarios were conducted based on data split ratios (70:30 and 80:20) with and without SMOTE, with performance evaluated using accuracy, precision, recall, and F1-score metrics. The results show that Logistic Regression with SMOTE achieved the best performance, with an accuracy of 83.80%, precision of 82.80%, recall of 82.70%, and F1-score of 82.80%. Sentiment analysis revealed a predominance of negative sentiment, reflecting public criticism of the coaching change. This study provides insights into public perception and demonstrates that Logistic Regression outperforms Naive Bayes in the context of social media sentiment analysis.

Keywords: Sentiment Analysis, Patrick Kluivert, Twitter/X, Naive Bayes, Logistic Regression, SMOTE

DAFTAR ISI

Halaman Pernyataan	ii
Halaman Pengesahan.....	iii
Halaman Persembahan.....	iv
Kata Pengantar	v
Abstrak.....	vi
Abstract.....	vii
Daftar Isi	viii
Daftar Tabel.....	x
Daftar Gambar	xi
Kode Program	xii
Daftar Lampiran	xiii
Bab I Pendahuluan.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Pembatasan Masalah	4
1.5 Sistematika Penulisan	4
Bab II Tinjauan Pustaka	6
2.1 Penelitian Terkait.....	6
2.2 Landasan Teori	7
2.2.1 Analisis Sentimen.....	7
2.2.2 Media Sosial	8
2.2.3 X (<i>Twitter</i>).....	8
2.2.4 Patrick Kluivert	9
2.2.5 Pengumpulan Data	10
2.2.6 <i>Text-Preprocessing</i>	11
2.2.7 Ekstraksi Fitur	12
2.2.8 SMOTE (<i>Synthetic Minority Over-sampling Technique</i>)	13
2.2.9 Klasifikasi.....	14
2.2.10 <i>Confusion Matrix</i>	17
Bab III Metodologi Penelitian	19
3.1 Pendekatan Penelitian.....	19
3.2 Perangkat yang Digunakan.....	19
3.2.1 Perangkat Keras.....	19
3.2.2 Perangkat Lunak.....	19
3.3 Langkah Penelitian	19
3.3.1 Pengumpulan Data	20

3.3.2	<i>Text-Preprocessing</i>	22
3.3.3	Ekstraksi Fitur	23
3.3.4	SMOTE	24
3.3.5	<i>Naive Bayes</i>	24
3.3.6	<i>Logistic Regression</i>	25
3.3.7	Evaluasi	26
3.3.8	Desain Skenario.....	27
Bab IV	Hasil dan Analisis.....	29
4.1	Pengumpulan Data.....	29
4.1.1	Metode Pengumpulan Data	29
4.1.2	Labeling.....	30
4.2	<i>Text-Preprocessing</i>	33
4.3	Hasil Ekstraksi Fitur	37
4.4	Hasil SMOTE	37
4.5	Implementasi Algoritma	39
4.5.1	<i>Naive Bayes</i>	39
4.5.2	<i>Logistic Regression</i>	39
4.6	Pengujian dan Evaluasi.....	40
4.6.1	Skenario Pengujian.....	41
4.6.2	Perbandingan Rata-Rata Skenario.....	44
4.7	Hasil Analisis.....	46
4.7.1	Kecenderungan Sentimen Publik	46
4.7.2	Skenario dengan Performa Paling Optimal	47
4.7.3	Perbandingan Hasil dengan Penelitian Terkait	49
Bab V	PENUTUP.....	52
5.1	Kesimpulan.....	52
5.2	Saran.....	52
Daftar Pustaka		53

DAFTAR TABEL

Tabel 2.1	<i>Confusion Matrix</i>	17
Tabel 3.1	Skenario Pengujian	27
Tabel 3.2	Rata-Rata Skenario Pengujian	28
Tabel 4.1	Pelabelan Sentimen Dengan Bantuan Empat AI	31
Tabel 4.2	Contoh <i>Case Folding</i>	33
Tabel 4.3	Contoh <i>Cleaning</i>	34
Tabel 4.4	Contoh <i>Normalize</i>	35
Tabel 4.5	Contoh <i>Tokenizing</i>	35
Tabel 4.6	Contoh <i>Stopword Removal</i>	36
Tabel 4.7	Contoh <i>Stemming</i>	36
Tabel 4.8	Skenario 1: <i>Naive Bayes</i> , 70:30, Tanpa SMOTE	41
Tabel 4.9	Skenario 2: <i>Naive Bayes</i> , 80:20, Tanpa SMOTE	41
Tabel 4.10	Skenario 3: <i>Naive Bayes</i> , 70:30, SMOTE	42
Tabel 4.11	Skenario 4: <i>Naive Bayes</i> , 80:20, SMOTE	42
Tabel 4.12	Skenario 5: <i>Logistic Regression</i> , 70:30, Tanpa SMOTE	42
Tabel 4.13	Skenario 6: <i>Logistic Regression</i> , 80:20, Tanpa SMOTE	43
Tabel 4.14	Skenario 7: <i>Logistic Regression</i> , 70:30, SMOTE	43
Tabel 4.15	Skenario 8: <i>Logistic Regression</i> , 80:20, SMOTE	43
Tabel 4.16	Hasil Pengujian Delapan Skenario	44
Tabel 4.17	Rata-Rata Skenario 7 dan 8 (<i>Naive Bayes</i> , Tanpa SMOTE)	45
Tabel 4.18	Rata-Rata Skenario 7 dan 8 (<i>Naive Bayes</i> , SMOTE)	45
Tabel 4.19	Rata-Rata Skenario 7 dan 8 (<i>Logistic Regression</i> , Tanpa SMOTE)	45
Tabel 4.20	Rata-Rata Skenario 7 dan 8 (<i>Logistic Regression</i> , SMOTE)	46

DAFTAR GAMBAR

Gambar 3.1	<i>Flowchart Analisis Sistem</i>	20
Gambar 3.2	<i>Flowchart Implementasi Algoritma Naive Bayes</i>	24
Gambar 3.3	<i>Flowchart Implementasi Algoritma Logistic Regression</i>	25
Gambar 4.1	Penggunaan AI dalam <i>Labeling</i>	30
Gambar 4.2	Jumlah Validasi Sentimen	32
Gambar 4.3	Distribusi Label Sentimen	32
Gambar 4.4	Hasil TF-IDF	37
Gambar 4.5	Hasil SMOTE pada <i>split</i> data 70:30	38
Gambar 4.6	Hasil SMOTE pada <i>split</i> data 80:20	38
Gambar 4.7	<i>Chart</i> Perbandingan Performa Rata-Rata Skenario	48

KODE PROGRAM

Kode Program 4.1	<i>Crawling</i>	30
Kode Program 4.2	<i>Case Folding.....</i>	33
Kode Program 4.3	<i>Cleaning.....</i>	34
Kode Program 4.4	<i>Normalize.....</i>	34
Kode Program 4.5	<i>Tokenizing.....</i>	35
Kode Program 4.6	<i>Stopword Removal</i>	36
Kode Program 4.7	<i>Stemming.....</i>	36
Kode Program 4.8	<i>TF-IDF</i>	37
Kode Program 4.9	<i>SMOTE.....</i>	38
Kode Program 4.10	<i>Naive Bayes</i>	39
Kode Program 4.11	<i>Logistic Regression.....</i>	39

DAFTAR LAMPIRAN

Lampiran A	Dataset Pergantian Pelatih Timnas	A-1
Lampiran B	Contoh Hasil <i>Case Folding</i>	B-1
Lampiran C	Contoh Hasil <i>Cleaning</i>	C-1
Lampiran D	Contoh Hasil <i>Normalize</i>	D-1
Lampiran E	Contoh Hasil <i>Tokenizing</i>	E-1
Lampiran F	Contoh Hasil <i>Stopword Removal</i>	F-1
Lampiran G	Contoh Hasil <i>Stemming</i>	G-1
Lampiran H	Contoh <i>Labeling</i>	H-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sepak bola merupakan olahraga paling populer di dunia yang tidak hanya berperan sebagai ajang kompetisi, tetapi juga sebagai sarana pemersatu lintas budaya, usia, dan latar belakang sosial[1]. Di Indonesia, sepak bola memiliki tempat istimewa di hati masyarakat, di mana Tim Nasional (Timnas) Indonesia menjadi simbol kebanggaan nasional dengan dukungan luar biasa dari para penggemar di seluruh penjuru negeri. Keberhasilan Timnas Indonesia lolos ke putaran ketiga Kualifikasi Piala Dunia 2026 menandai tonggak penting dalam perkembangan sepak bola nasional dan menunjukkan peningkatan performa tim yang signifikan. Namun, di tengah euforia tersebut, Persatuan Sepak Bola Seluruh Indonesia (PSSI) membuat keputusan kontroversial pada Januari 2025 dengan mengganti pelatih kepala Timnas Indonesia, meskipun pelatih sebelumnya dinilai telah membawa dampak positif dan konsisten sejak 2020. Keputusan ini menimbulkan beragam tanggapan dari publik, baik berupa kritik maupun dukungan, yang disampaikan secara masif melalui media sosial, khususnya platform X (sebelumnya *Twitter*).

Fenomena ini menunjukkan bagaimana keputusan strategis dalam dunia olahraga, khususnya yang melibatkan Timnas Indonesia, tidak hanya berdampak secara teknis, tetapi juga memicu dinamika opini publik. Media sosial menjadi medium utama bagi masyarakat untuk mengekspresikan pandangan mereka, sehingga menghasilkan beragam sentimen yang mencerminkan penerimaan maupun penolakan terhadap kebijakan PSSI.

Dalam konteks ini, analisis sentimen menjadi pendekatan yang relevan untuk memahami persepsi masyarakat terhadap pergantian pelatih Timnas Indonesia. Dengan menerapkan teknik klasifikasi, opini publik yang terekam dalam media sosial dapat dikelompokkan ke dalam beberapa kategori sentimen, seperti positif, negatif, dan netral. Klasifikasi ini memungkinkan peneliti untuk mengidentifikasi kecenderungan umum respon masyarakat—apakah mendukung

atau menolak terhadap kebijakan tersebut. Beragam algoritma klasifikasi dapat diterapkan dalam analisis sentimen, di antaranya adalah *Naive Bayes* dan *Logistic Regression*. *Naive Bayes* merupakan salah satu algoritma dalam pembelajaran mesin yang berlandaskan pada prinsip *Teorema Bayes*, dengan asumsi dasar bahwa setiap atribut atau fitur pada data bersifat independen secara kondisional terhadap kelas target[2], sedangkan *Logistic Regression* adalah metode klasifikasi yang digunakan untuk membangun model hubungan antara variabel independen dan variabel dependen biner, dengan cara memetakan *input* berupa variabel diskrit maupun kontinu ke dalam probabilitas *output* yang terbagi ke dalam dua kelas, seperti positif atau negatif[3].

Data yang diperoleh dalam penelitian ini memiliki distribusi kelas yang tidak seimbang, sehingga metode SMOTE (*Synthetic Minority Over-sampling Technique*) digunakan untuk menyeimbangkan data tersebut. SMOTE merupakan teknik yang digunakan untuk menyeimbangkan distribusi data dengan cara menghasilkan sampel sintesis pada kelas minoritas hingga jumlahnya setara dengan kelas mayoritas[4]. *Confusion Matrix* adalah metode evaluasi kinerja model klasifikasi yang digunakan untuk mengidentifikasi ketepatan prediksi terhadap data. Matriks ini menyajikan perbandingan antara label hasil prediksi dan label aktual, yang mencerminkan sejauh mana model berhasil mengklasifikasikan data dengan benar atau keliru[5].

Penelitian terkait analisis sentimen di media sosial telah banyak dilakukan. Sebagai contoh, penelitian yang menganalisis sentimen masyarakat Indonesia terhadap perkembangan teknologi *metaverse* di *platform X* menggunakan *Naive Bayes* dan *Logistic Regression* menemukan bahwa *Logistic Regression* dengan optimasi SMOTE menghasilkan akurasi tertinggi (95%) dibandingkan *Naive Bayes* (91%), dengan peningkatan signifikan pada *precision*, *recall*, dan *F1-score* untuk sentimen negatif setelah penerapan SMOTE[3]. Temuan ini menunjukkan bahwa *Logistic Regression* lebih unggul dalam menangani data tidak seimbang, yang relevan dengan konteks penelitian ini, mengingat sentimen negatif terhadap pergantian pelatih cenderung lebih sedikit dibandingkan sentimen positif atau netral. Wawasan dari penelitian tersebut

memperkuat pentingnya komparasi algoritma dan optimasi data untuk memastikan hasil analisis sentimen yang akurat dan representatif.

Selain itu, penelitian yang membandingkan kinerja *Multinomial Naive Bayes* dan *Logistic Regression* dalam analisis sentimen ulasan film pada *dataset* IMDB, yang terdiri dari 40.000 ulasan pengguna, menggunakan teknik pra-pemrosesan seperti *case folding*, *tokenizing*, *stemming*, *stopwords removal*, dan TF-IDF untuk mempersiapkan data teks, diikuti dengan klasifikasi menggunakan kedua algoritma tersebut. Hasilnya menunjukkan bahwa *Multinomial Naive Bayes* mencapai akurasi sebesar 85%, sedikit lebih rendah dibandingkan *Logistic Regression*[6]. Penelitian ini menegaskan bahwa *Multinomial Naive Bayes* sangat efektif untuk klasifikasi teks dengan data berdimensi tinggi, seperti ulasan film, yang memiliki kemiripan dengan data teks pendek di media sosial X. Temuan ini relevan untuk penelitian ini karena menunjukkan potensi *Multinomial Naive Bayes* dalam menangani data teks singkat seperti cuitan, sekaligus memperkuat perlunya komparasi dengan *Logistic Regression* untuk menentukan algoritma yang optimal.

Berdasarkan pendekatan serupa, penelitian ini bertujuan untuk mengetahui hasil analisis sentimen yang dapat memberikan wawasan bagi pengambil kebijakan di federasi sepak bola, pelatih, dan pemangku kepentingan lainnya untuk memahami penerimaan publik terhadap keputusan strategis. Selain itu, komparasi performa algoritma *Naive Bayes* dan *Logistic Regression* akan memberikan gambaran tentang algoritma mana yang lebih optimal dalam konteks analisis sentimen berbasis data media sosial. Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan metode analisis sentimen tetapi juga pada pengambilan kebijakan yang lebih tepat di dunia olahraga Indonesia.

1.2 Perumusan Masalah

Berdasarkan pada latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah bagaimana perbandingan performa algoritma *Naive Bayes* dan *Logistic Regression* dalam menganalisis sentimen publik terhadap pergantian pelatih Timnas Indonesia di media sosial X (*Twitter*)?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah untuk mengetahui hasil perbandingan performa algoritma *Naive Bayes* dan *Logistic Regression* dalam menganalisis sentimen publik terhadap pergantian pelatih Timnas Indonesia di media sosial X (*Twitter*)?

1.4 Pembatasan Masalah

Batasan masalah yang terdapat pada penelitian ini sebagai berikut:

- a. Sumber data yang digunakan adalah data *tweet* dari X (*Twitter*) yang diambil dari tanggal 6 Januari 2025 sampai 6 Maret 2025.
- b. Penggunaan *tweet* berbahasa Indonesia.
- c. Pengambilan data dengan cara *crawling* menggunakan *Google Colab*.
- d. Pelabelan dilakukan dengan bantuan empat *platform* AI, yaitu *ChatGPT*, *Copilot*, *Grok*, dan *Blackbox AI*.
- e. Kelas sentimen terdiri dari positif, negatif dan netral.

1.5 Sistematika Penulisan

Sistematika dari penulisan skripsi ini terdiri dari lima bab yaitu, Bab I Pendahuluan, Bab II Tinjauan Pustaka, Bab III Metodologi Penelitian, Bab IV Hasil dan Analisis, dan Bab V Penutup.

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, identifikasi masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas mengenai dasar teori yang mendukung penelitian mengenai analisis sentimen terhadap pergantian pelatih Timnas Indonesia pada fase kualifikasi piala dunia 2026 menggunakan algoritma *Naive Bayes* dan *Logistic Regression* pada media sosial X (*Twitter*).

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tentang metode yang digunakan, yaitu metode pengumpulan data dan metode untuk analisis data. Untuk analisis data menggunakan algoritma *Naive Bayes* dan *Logistic Regression*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menggambarkan profil singkat objek penelitian dan membahas hasil analisis sentimen terhadap pergantian pelatih Timnas Indonesia pada fase kualifikasi piala dunia 2026 berdasarkan penelitian yang sudah dilakukan dengan algoritma yang telah ditentukan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian berdasarkan rumusan masalah terkait analisis sentimen terhadap pergantian pelatih Timnas Indonesia pada fase kualifikasi piala dunia 2026 serta berisi saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.

DAFTAR PUSTAKA