

**Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia
menggunakan Algoritma *Multinomial Naïve Bayes*
dengan Optimasi *Particle Swarm Optimization***

SKRIPSI

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



disusun oleh
ICHLASUL AMAL WIJAKSANA
17.11.1744

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

**Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia
menggunakan Algoritma *Multinomial Naïve Bayes*
dengan Optimasi *Particle Swarm Optimization***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
ICHLASUL AMAL WIJAKSANA
17.11.1744

Kepada
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

HALAMAN PERSETUJUAN

SKRIPSI

**Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia
menggunakan Algoritma *Multinomial Naïve Bayes*
dengan Optimasi *Particle Swarm Optimization***

yang disusun dan diajukan oleh

**Ichlasul Amal Wijaksana
17.11.1744**

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 26 Juni 2023

Dosen Pembimbing,


**Anggit Dwi Hartanto, M.Kom
NIK. 190302163**

HALAMAN PENGESAHAN

SKRIPSI

**Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia
menggunakan Algoritma *Multinomial Naïve Bayes*
dengan Optimasi *Particle Swarm Optimization***

yang disusun dan diajukan oleh

ICHLASUL AMAL WIJAKSANA

17.11.1744

Telah dipertahankan di depan Dewan Penguji
pada tanggal 26 Juni 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Arifivanto Hadinegoro, S.Kom, MT
NIK. 190302289



Windha Mega Pradnya D, M.Kom
NIK. 190302185



Majid Rahardi, S.Kom., M.Eng
NIK. 190302393



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 26 Juni 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Farid, S.Kom, M.Kom,
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini,

Nama mahasiswa : ICHLASUL AMAL WIJAKSANA
NIM : 17.11.1744

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia menggunakan Algoritma *Multinomial Naïve Bayes* dengan Optimasi *Particle Swarm Optimization*

Dosen Pembimbing : Anggit Dwi Hartanto, 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 ketidak benaran 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, 26 Juni 2023

Yang Menyatakan,



ICHLASUL AMAL WIJAKSANA

HALAMAN PERSEMBAHAN

Dengan rasa syukur sebanyak-banyaknya skripsi ini dipersembahkan:

1. Keluarga penulis yang selalu memberikan dukungan dan bantuan dalam proses penyelesaian skripsi ini.
2. Civitas akademika Universitas Amikom Yogyakarta, termasuk dosen, karyawan, dan mahasiswa lainnya yang telah memberikan bantuan dan dukungan yang berarti selama penulisan skripsi ini.
3. Teman-teman penulis, baik dari lingkungan perkuliahan di Universitas Amikom Yogyakarta maupun dari universitas lain, yang telah memberikan bantuan, semangat, dan motivasi yang luar biasa dalam perjalanan penyelesaian skripsi ini.
4. Teman-teman penulis dari lingkungan sepeneliti rumah, yang juga telah memberikan dukungan, semangat, dan motivasi yang tak ternilai harganya dalam proses penulisan skripsi ini.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan karunia-Nya, sehingga penulis dapat menyelesaikan Skripsi yang berjudul “Analisis Sentimen Opini Publik tentang Isu LGBT di Indonesia menggunakan Algoritma Multinomial Naïve Bayes dengan Optimasi Particle Swarm Optimization” dengan lancar. Skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Program Studi Informatika di Universitas Amikom Yogyakarta.

Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan bantuan dalam penyelesaian skripsi ini. Khususnya, penulis mengucapkan terima kasih kepada:

- Bapak Hanif Al Fatta, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta, atas bimbingan dan kesempatan yang diberikan kepada penulis selama proses penulisan skripsi.
- Bapak Anggit Dwi Hartanto, M.Kom., selaku pembimbing penulis, atas arahan, bimbingan, dan ilmu yang telah beliau bagikan sehingga skripsi ini dapat diselesaikan.
- Para dosen Program Studi Informatika Amikom Yogyakarta yang telah memberikan ilmu pengetahuan dan pengalaman berharga selama penulis menempuh perkuliahan.

Penulis juga ingin berterima kasih kepada semua pihak yang tidak dapat disebutkan satu per satu, yang turut berperan serta memberikan dukungan dan motivasi dalam perjalanan penulisan skripsi ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini mungkin masih terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran konstruktif dari pembaca untuk perbaikan di masa yang akan datang.

Akhir kata, semoga laporan skripsi ini dapat memberikan manfaat dan kontribusi positif bagi pengembangan ilmu pengetahuan dan teknologi. Terima kasih atas perhatian dan dukungan dari semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Yogyakarta, 26 Juni 2023

Penulis

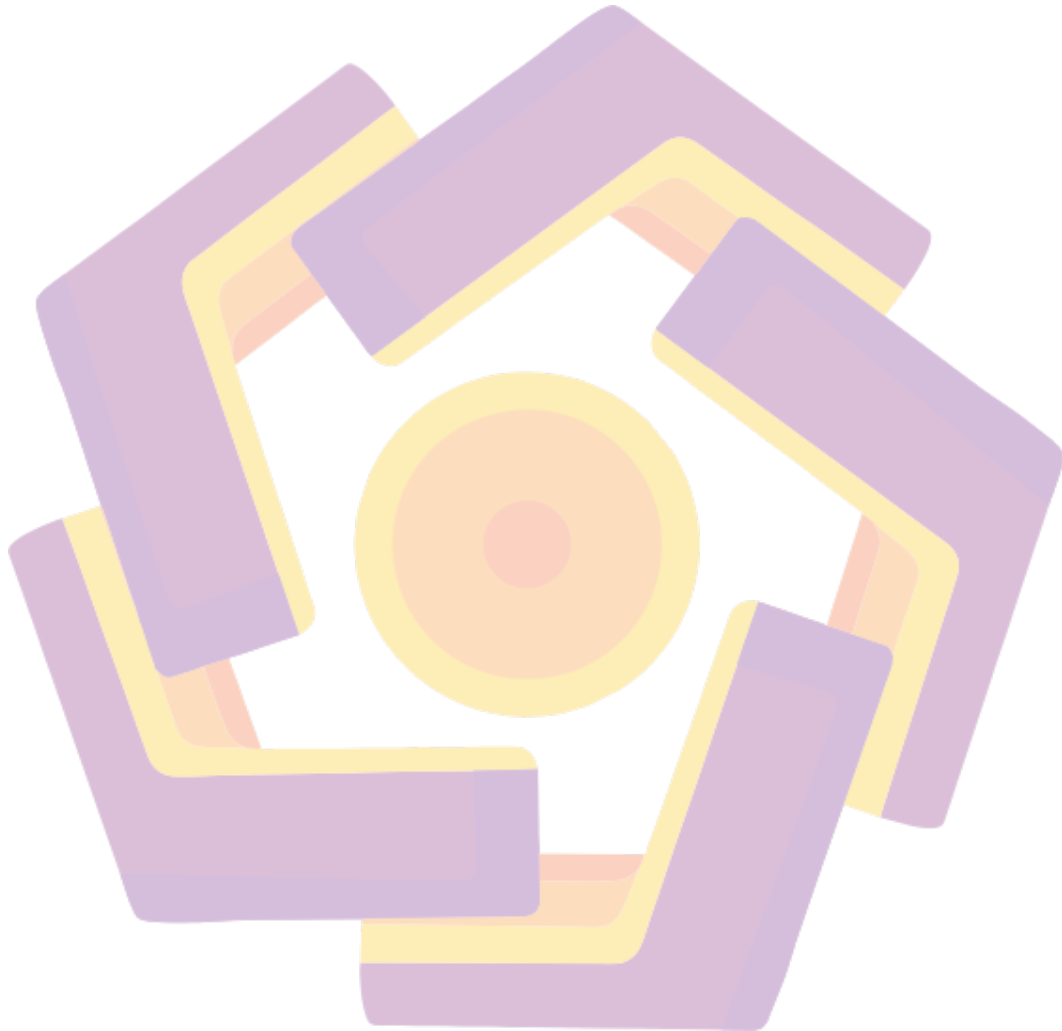
DAFTAR ISI

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
DAFTAR LAMPIRAN.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	11
2.2.1 Online Review	11
2.2.2 Analisis Sentimen	11
2.2.3 Data Mining	11
2.2.4 Preprocessing.....	12
2.2.5 Ekstraksi Fitur dan Pembobotan (TF-IDF)	13

2.2.6 Vader sentiment	13
2.2.7 Klasifikasi	14
2.2.8 Machine learning	15
2.2.9 Klasifikasi Multinomial Naïve Bayes	16
2.2.10 PSO (Particle Swarm Optimization)	16
BAB III METODE PENELITIAN	19
3.1 Alur Penelitian	19
3.2 Metode Analisis	20
3.3 Alat dan Bahan	23
3.3.1 Variabel Penelitian	23
3.3.2 Bahan Penelitian	24
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Pengambilan Data	25
4.2 Preprocessing data	28
4.2.1 Preprocessing data tweet dalam bahasa indonesia	30
4.1.2 Translate Data ke Bahasa Inggris	32
4.1.2 Preprocessing ulang data yang sudah di translate	33
4.1.3 Label menggunakan Vader sentiment	35
4.3 Klasifikasi Multinomial Naïve Bayes	37
4.4 Klasifikasi Multinomial Naïve Bayes Optimasi dengan PSO	39
4.4 Analisis Hasil	45
4.4.1 Hasil Klasifikasi Multinomial Naïve Bayes	45
4.4.2 Hasil Klasifikasi Multinomial Naïve Bayes Optimasi dengan PSO	46
4.4.3 Hasil Sentimen Distribusi	48
4.4.1 Word Cloud	49
4.4.3 Pembahasan Perbandingan	51
BAB V PENUTUP	52
5.1 Kesimpulan	52
5.2 Saran	52
REFERENSI	54
LAMPIRAN	56

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian.....	8
Tabel 3. 1 Sekenario Pengujian.....	23
Tabel 4. 1 Perbandingan Evaluasi MNB & MNB+PSO.....	51

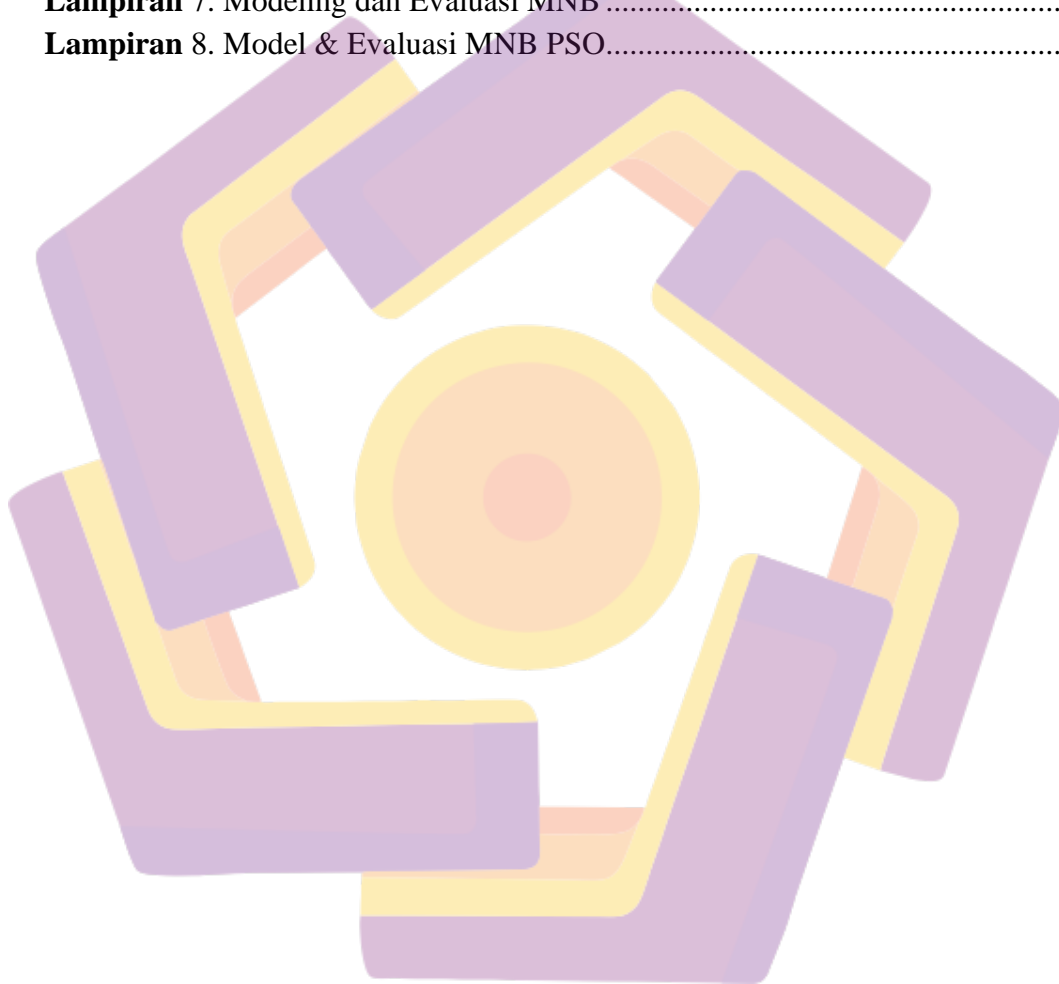


DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	19
Gambar 4. 1 <i>script scrapping</i> data	25
Gambar 4. 2 <i>script scrapping</i> data	25
Gambar 4. 3 penggabungan kolom	28
Gambar 4. 4 Preprocessing bahasa Indonesia	30
Gambar 4. 5 Preprocessing bahasa Indonesia	31
Gambar 4. 6 Hasil preproceesing	32
Gambar 4. 7 Terjemah ke bahasa Inggris	32
Gambar 4. 8 Hasil terjemahan.....	33
Gambar 4. 9 <i>Preprocessing</i> bahasa Inggris	33
Gambar 4. 10 pelabelan menggunakan <i>vader sentiment</i>	35
Gambar 4. 11 hasil pelabelan	36
Gambar 4. 12 Klasifikasi <i>Multinomial Naïve Bayes</i>	37
Gambar 4. 13 baca file dan pemilihan target x dan y	39
Gambar 4. 14 <i>Vektorisasi, Smote</i> dan <i>Split</i> data	39
Gambar 4. 15 Sleksi fitur PSO	40
Gambar 4. 16 Penerapan model dan uji model	43
Gambar 4. 17 Visualisasi <i>Wordcloud</i>	44
Gambar 4. 18 <i>Confusion Matrix & Evaluasi MNB</i>	45
Gambar 4. 19 Hasil seleksi fitur	46
Gambar 4. 20 <i>Confusion Matrix & Evaluasi MNB + PSO</i>	47
Gambar 4. 21 Sentimen distribusi	49
Gambar 4. 22 <i>Wordcloud</i> Positif	49
Gambar 4. 23 <i>Wordcloud</i> Negatif	50
Gambar 4. 24 <i>Wordcloud</i> Netral	50

DAFTAR LAMPIRAN

Lampiran 1. script pengambilan data	56
Lampiran 2. Script gabung kolom dan baris	57
Lampiran 3. <i>Preprocessing</i> dalam bahasa Indonesia	57
Lampiran 4. Terjemah ke bahasa Inggris import pandas as pd.....	60
Lampiran 5. <i>Preprocessing</i> dalam bahasa Inggris	61
Lampiran 6. <i>Lebelling</i> menggunakan <i>Vader sentiment</i>	62
Lampiran 7. Modeling dan Evaluasi MNB	63
Lampiran 8. Model & Evaluasi MNB PSO.....	65




DAFTAR LAMBANG DAN SINGKATAN

gbest	nilai terbaik global
MNB	<i>Multinomial Naïve Bayes</i>
NLTK	<i>Natural Language Toolkit</i>
pbest _i	nilai terbaik partikel i
PSO	<i>Particle Swarm Optimization</i>
SMOTE	<i>Synthetic Minority Oversampling Technique</i>
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>
VADER Sentiment	<i>Valence Aware Dictionary and sEntiment Reasoner</i>



DAFTAR ISTILAH



Analisis sentimen	pengenalan dan ekstraksi opini dari teks.
Isu LGBT	isu tentang lesbian, gay, biseksual, dan transgender.
<i>Multinomial Naïve Bayes</i> (MNB)	algoritma klasifikasi berbasis probabilitas.
<i>Particle Swarm Optimization</i> (PSO)	algoritma optimasi berbasis gerombolan.
Twitter	media sosial berbasis <i>tweet</i> .
<i>Preprocessing</i>	pembersihan dan transformasi data mentah.
NLTK	modul Python untuk pemrosesan bahasa alami.
VADER Sentiment	algoritma analisis sentimen untuk teks media sosial.
TF-IDF	metode pembobotan kata berdasarkan frekuensi.
SMOTE	metode <i>oversampling</i> dengan sampel sintetis.

INTISARI

Isu LGBT (lesbian, gay, biseksual, dan transgender) di Indonesia menimbulkan pro dan kontra di masyarakat. Media sosial Twitter digunakan oleh masyarakat untuk menyuarakan opini dan sikap mereka terhadap isu tersebut. Penelitian ini membandingkan kinerja algoritma *Multinomial Naïve Bayes* (MNB) dan MNB dengan fitur seleksi *Particle Swarm Optimization* (PSO) dalam mengklasifikasikan sentimen opini publik. Penelitian ini menggunakan data *tweet* yang mengandung kata kunci lgbt. Data *tweet* tersebut dilakukan *preprocessing* dengan bahasa Indonesia, diterjemahkan ke bahasa Inggris, dan *preprocessing* lagi menggunakan NLTK. Selanjutnya, data *tweet* tersebut dilakukan pelabelan sentimen menggunakan VADER Sentiment.

Kemudian data *tweet* tersebut diklasifikasikan menggunakan algoritma MNB dan MNB dengan fitur seleksi PSO. Sebelum klasifikasi, data *tweet* tersebut dilakukan pembobotan TF-IDF, *oversampling* dengan SMOTE, dan pemisahan data menjadi 80% data latih dan 20% data uji. Hasil penelitian menunjukkan bahwa algoritma MNB dengan fitur seleksi PSO memiliki akurasi dan performa yang lebih baik daripada algoritma MNB standar dalam mengklasifikasikan sentimen opini publik tentang isu LGBT di Indonesia. Algoritma MNB dengan fitur seleksi PSO menghasilkan akurasi sebesar 0.98, sedangkan algoritma MNB standar menghasilkan akurasi sebesar 0.97. Algoritma MNB dengan fitur seleksi PSO juga menunjukkan peningkatan dalam nilai *precision*, *recall*, dan *F1-score*.

Kata kunci: analisis sentimen, isu LGBT, *Multinomial Naïve Bayes*, *Particle Swarm Optimization*, Twitter

ABSTRACT

The issue of LGBT (lesbian, gay, bisexual, and transgender) in Indonesia causes pros and cons in society. Social media Twitter is used by people to express their opinions and attitudes towards the issue. This study compares the performance of the Multinomial Naïve Bayes (MNB) algorithm and MNB with Particle Swarm Optimization (PSO) feature selection in classifying public opinion sentiment. This study uses tweet data that contains the keyword lgbt. The tweet data is preprocessed with Indonesian, translated into English, and preprocessed again using NLTK. Furthermore, the tweet data is labeled sentiment using VADER Sentiment.

Then the tweet data is classified using the MNB algorithm and MNB with PSO feature selection. Before classification, the tweet data is weighted by TF-IDF, oversampled with SMOTE, and split into 80% training data and 20% testing data. The results show that the MNB algorithm with PSO feature selection has better accuracy and performance than the standard MNB algorithm in classifying public opinion sentiment about the LGBT issue in Indonesia. The MNB algorithm with PSO feature selection produces an accuracy of 0.98, while the standard MNB algorithm produces an accuracy of 0.97. The MNB algorithm with PSO feature selection also shows an increase in precision, recall, and F1-score values.

Keyword: *Sentiment Analysis, LGBT issue, Multinomial Naïve Bayes, Particle Swarm Optimization, Twitter*