

**OPTIMASI HYPERPARAMETER PADA ALGORITMA
NAIVE BAYES DAN RANDOM FOREST UNTUK ANALISIS
SENTIMEN APLIKASI RILIV**

SKRIPSI

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



disusun oleh

KRISNA EKA MAULANA

21.11.4249

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**OPTIMASI HYPERPARAMETER PADA ALGORITMA
NAIVE BAYES DAN RANDOM FOREST UNTUK ANALISIS
SENTIMEN APLIKASI RILIV**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

KRISNA EKA MAULANA

21.11.4249

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**OPTIMASI HYPERPARAMETER PADA ALGORITMA NAIVE BAYES
DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN APLIKASI**

RILIV

yang disusun dan diajukan oleh

Krisna Eka Maulana

21.11.4249

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 29 Juli 2025

Dosen Pembimbing,



Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom.

NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

**OPTIMASI HYPERPARAMETER PADA ALGORITMA NAIVE BAYES
DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN APLIKASI
RILIV**

yang disusun dan diajukan oleh

Krisna Eka Maulana

21.11.4249

Telah dipertahankan di depan Dewan Penguji
pada tanggal 29 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302

Ali Mustopa, S.Kom., M.Kom
NIK. 190302192

Windha Mega Pradnya D, S.Kom., M.Kom
NIK. 190302185

Tanda Tangan



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 29 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Krisna Eka Maulana
NIM : 21.11.4249

Menyatakan bahwa Skripsi dengan judul berikut:

OPTIMASI HYPERPARAMETER PADA ALGORITMA NAIVE BAYES DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN APLIKASI RILIV

Dosen Pembimbing : Windha Mega Pradnya Dhuhiita, 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, 29 Juli 2025

Yang Menyatakan,



Krisna Eka Maulana

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam atas kehadiran Allah SWT, yang telah memberikan kekuatan, kesehatan, dan kelapangan waktu, sehingga skripsi ini dapat terselesaikan dengan baik, penulis persembahkan skripsi ini kepada:

1. Mama Kasmianti dan Ayah Sugeng tercinta, terima kasih atas doa yang tiada henti, kasih sayang yang tulus, pengorbanan yang tak tergantikan dan dukungan serta bimbingan dalam setiap langkah. Semoga mama dan ayah senantiasa sehat, panjang umur, lancar rezeki dan bahagia selalu.
2. Adik tercinta, yang senantiasa memberikan dukungan dan semangat. Terima kasih telah menjadi penyemangat.
3. Keluarga besar tersayang, Tante sukut, Om Ali, Icha yang senantiasa memberikan dukungan dan doa restu dalam setiap langkah perjalanan ini.
4. Pasangan tersayang, Eka Agustin Yulistiyah yang senantiasa menemani penulis dalam menyelesaikan skripsi ini. Kehadiranmu menjadikan sumber kebahagiaan dan inspirasi dalam setiap langkah. Terima kasih atas dukungan, kesabaran dan waktu yang telah diberikan dengan tulus.
5. Teman-teman Himpunan Mahasiswa Informatika, terima kasih atas waktu kebersamaan, dukungan dan pengalaman yang sangat penuh makna.
6. Sahabat-sahabat kelas dan kontrakan penulis, terima kasih telah menemani dan dukung penulis.
7. Terima kasih kepada diriku sendiri yang mampu selalu bertahan dalam kondisi apapun, sehingga dapat menyelesaikan skripsi ini dengan baik.
8. Almamater tercinta, Universitas Amikom Yogyakarta, tempat penulis dalam menimba ilmu, pembentukan karakter, dan mengembangkan potensi diri.

Skripsi ini adalah bukti atas ketekunan, doa, dan semangat pantang menyerah akan selalu menemukan jalan. Semoga skripsi ini dapat memberikan manfaat, baik bagi penulis maupun bagi pembaca di masa yang akan datang.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi berjudul **“OPTIMASI HYPERPARAMETER PADA ALGORITMA NAIVE BAYES DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN APLIKASI RILIV”**. Skripsi ini disusun sebagai salah satu syarat untuk meraih gelar sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, mengingat keterbatasan kemampuan, pengetahuan, dan pengalaman saya. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun agar karya tulis berikutnya dapat lebih baik lagi.

Proses penyusunan skripsi ini tidak lepas dari berbagai tantangan, namun berkat bantuan, bimbingan, dan kerja sama dari banyak pihak, baik secara moril maupun materil, penulis dapat menyelesaikannya dengan baik. Oleh karena itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. M. Suyanto, MM., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr. Kusrini, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom, selaku Ketua Program Studi Informatika Universitas Amikom Yogyakarta.
4. Ibu Windha Mega Pradnya Dhuhiya, S.Kom., M.Kom., selaku dosen pembimbing yang dengan sabar dan ikhlas meluangkan waktu, tenaga dan pikiran, serta memberikan bimbingan arahan, motivasi dan saran-saran yang sangat berharga kepada penulis selama menyusun skripsi ini.
5. Bapak Nuri Cahyono, M.Kom, selaku dosen wali yang memberikan bimbingan dan bantuan selama penulis menempuh masa studi di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

6. Seluruh staff dosen Universitas Amikom Yogyakarta, khususnya Fakultas Ilmu Komputer Program Studi Informatika yang telah memberikan ilmu yang tidak ternilai harganya, masukan, pemikiran dan tenaga selama proses pembelajaran yang dapat menambah wawasan bagi penulis.
7. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan dukungan kepada penulis.

Semoga Allah SWT memberikan balasan dengan berlipat ganda kepada semuanya yang telah membantu dalam penulisan skripsi ini. Penulis berharap skripsi ini dapat memberikan manfaat khususnya bagi penulis dan para pembaca.

Yogyakarta, 26 Juli 2025

Penulis

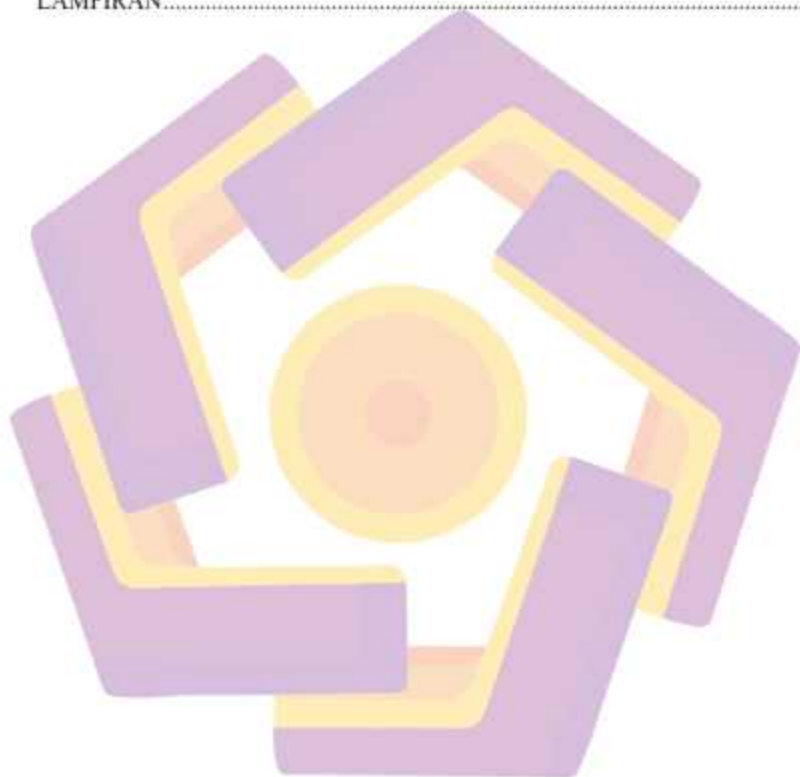


DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiv
DAFTAR LAMBANG DAN SINGKATAN	xv
DAFTAR ISTILAH	xvi
INTISARI	xvii
ABSTRACT	xviii
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
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	12
2.2.1 Riliv	12
2.2.2 Analisis Sentimen	12
2.2.3 Data Scraping	13
2.2.4 Pra-pemrosesan Teks	13
2.2.5 TF-IDF (<i>Term Frequency - Inverse Document Frequency</i>)	14

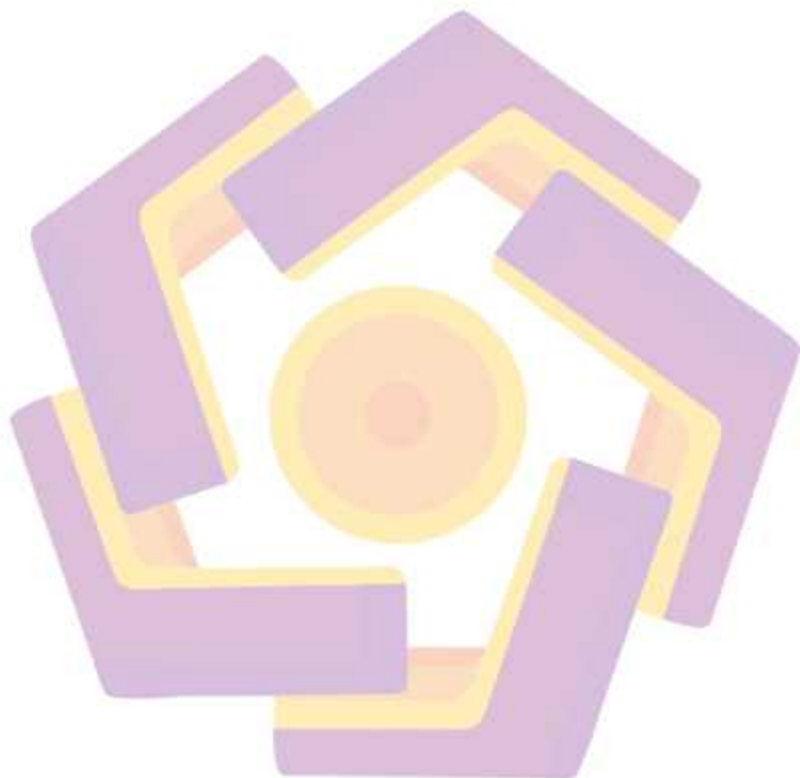
2.2.6	SMOTE	15
2.2.7	Naive Bayes	15
2.2.8	Random Forest	16
2.2.9	<i>Optimazer dan Hyperparameter Tuning Optuna</i>	17
2.2.10	Evaluasi Model	18
BAB III METODE PENELITIAN		20
3.1	Objek Penelitian	20
3.2	Alur Penelitian	20
3.2.1	Pengumpulan Data	21
3.2.2	Pelabelan Data	21
3.2.3	<i>Preprocessing Data</i>	21
3.2.4	TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>)	23
3.2.5	SMOTE <i>Oversampling</i>	23
3.2.6	<i>Modeling Random Forest</i>	23
3.2.7	<i>Modeling Naive Bayes</i>	24
3.2.8	<i>Modeling dan Training Optuna</i>	24
3.2.9	Evaluasi Model	25
3.3	Alat dan Bahan	25
3.3.1	Data Penelitian	25
3.3.2	Alat/Instrumen	25
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Deskripsi Dataset	27
4.2	Pelabelan Data	27
4.3	<i>Preprocessing Data</i>	28
4.4	Pembobotan TF-IDF	38
4.5	Split Data	38
4.6	SMOTE	38
4.7	<i>Modeling Random Forest</i>	39
4.8	<i>Modeling Naive Bayes</i>	40
4.9	Evaluasi Model	42
4.10	<i>Modeling dan Training Optuna</i>	44

4.11 Deploy.....	46
BAB V PENUTUP	48
5.1 Kesimpulan.....	48
5.2 Saran.....	48
REFERENSI	50
LAMPIRAN.....	54



DAFTAR TABEL

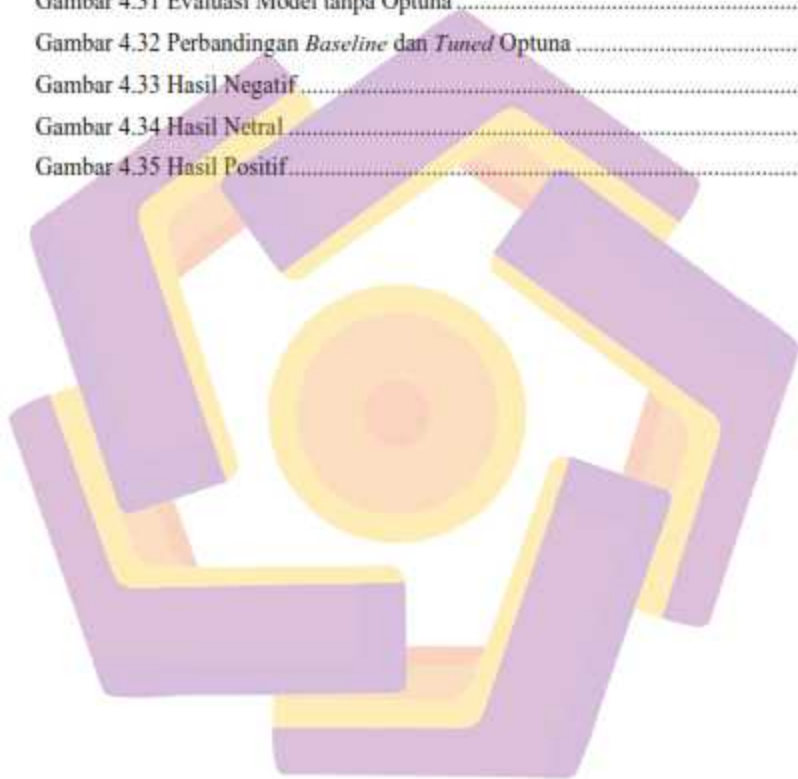
Tabel 2.1 Keaslian Penelitian	7
-------------------------------------	---



DAFTAR GAMBAR

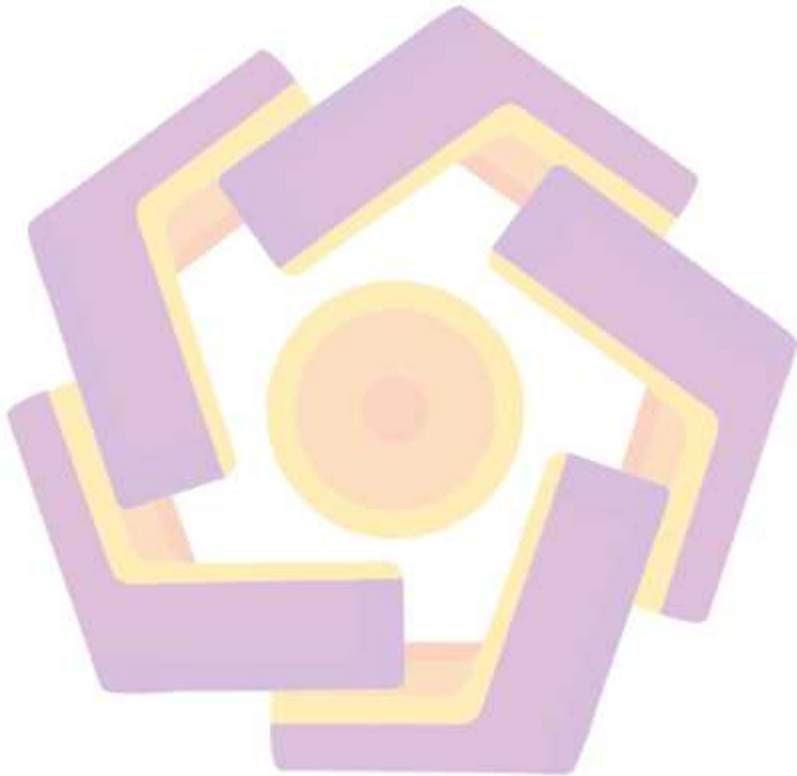
Gambar 2.1 Logo Riliv	12
Gambar 2.2 Proses Prediksi Random Forest.....	17
Gambar 2.3 Prosedur Dasar Optuna untuk Penyesuaian <i>Hyperparameter</i>	18
Gambar 3.1 Alur Penelitian	20
Gambar 4.1 Deskripsi Dataset	27
Gambar 4.2 Hasil Pelabelan Data	28
Gambar 4.3 <i>Source Code Case Folding</i>	28
Gambar 4.4 Hasil dari <i>Case Folding</i>	29
Gambar 4.5 <i>Source Code</i> dari <i>Remove Non-Alphabet</i>	30
Gambar 4.6 Hasil dari <i>Remove Non-Alphabet</i>	30
Gambar 4.7 <i>Source Code Translate Emoji</i>	30
Gambar 4.8 Hasil dari <i>Translate Emoji</i>	31
Gambar 4.9 <i>Source Code</i> dari <i>Remove Punctuation</i>	31
Gambar 4.10 Hasil dari <i>Remove Punctuation</i>	32
Gambar 4.11 <i>Source Code</i> dari <i>Replace Slang</i>	32
Gambar 4.12 Hasil dari <i>Replace Slang</i>	33
Gambar 4.13 <i>Source Code</i> dari <i>Replace Word Elongation</i>	33
Gambar 4.14 Hasil dari <i>Replace Word Elongation</i>	33
Gambar 4.15 <i>Source Code Tokenizing</i>	34
Gambar 4.16 Hasil <i>Tokenizing</i>	35
Gambar 4.17 <i>Source Code Stopword Removal</i>	36
Gambar 4.18 Hasil <i>Stopword Removal</i>	36
Gambar 4.19 <i>Source Code Stemming</i>	37
Gambar 4.20 Hasil <i>Stemming</i>	37
Gambar 4.21 Hasil TF-IDF.....	38
Gambar 4.22 Hasil SMOTE.....	39
Gambar 4.23 Implementasi Random Forest	40
Gambar 4.24 Hasil Random Forest.....	40
Gambar 4.25 Implementasi Naive Bayes.....	41

Gambar 4.26 Hasil Naive Bayes	41
Gambar 4.27 <i>Confusion Matrix</i> Naive Bayes dan Random Forest.....	42
Gambar 2.28 Perbandingan <i>Precision</i> , <i>Recall</i> , dan <i>F1-Score</i> tiap Sentimen	43
Gambar 2.29 Perbandingan Akurasi Model.....	43
Gambar 4.30 Evaluasi Model dengan Optuna	44
Gambar 4.31 Evaluasi Model tanpa Optuna	45
Gambar 4.32 Perbandingan <i>Baseline</i> dan <i>Tuned Optuna</i>	46
Gambar 4.33 Hasil Negatif.....	46
Gambar 4.34 Hasil Netral.....	47
Gambar 4.35 Hasil Positif.....	47



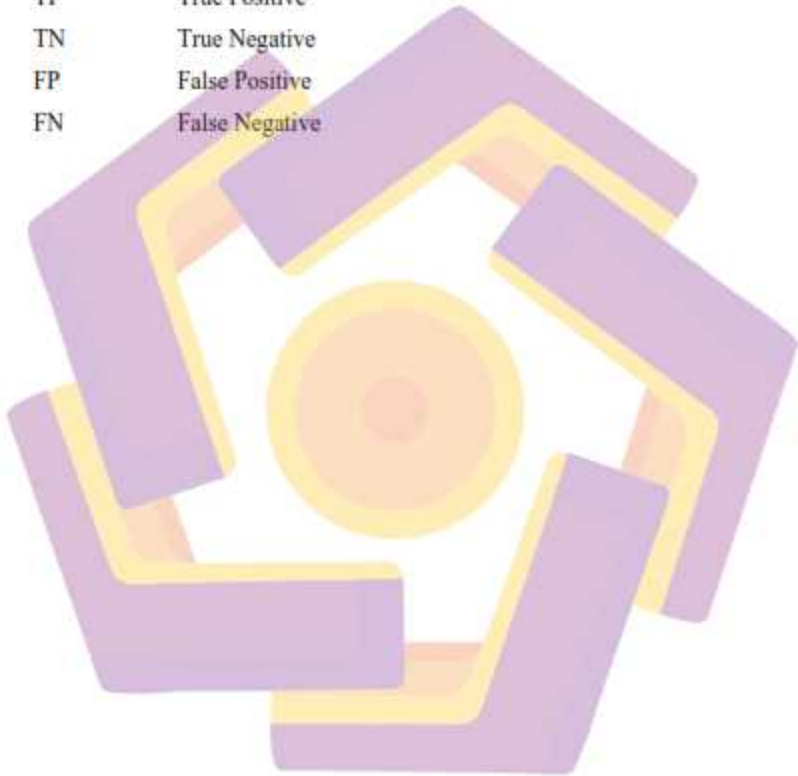
DAFTAR LAMPIRAN

Lampiran 1. Link Dataset Riliv	54
Lampiran 2. Link Model Random Forest dan Naive Bayes	54
Lampiran 3. Link Konfigurasi Streamlit.....	54



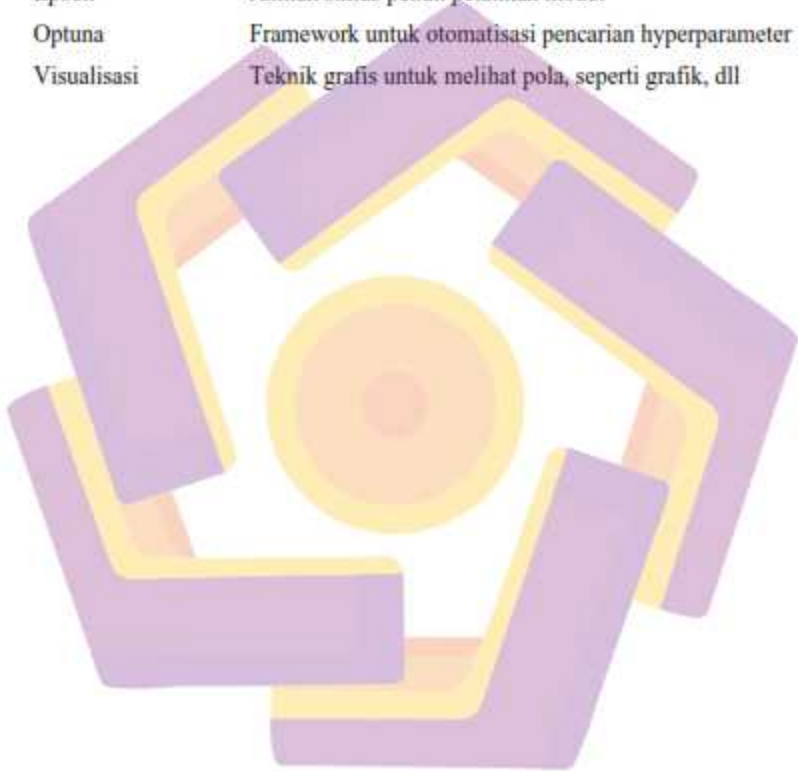
DAFTAR LAMBANG DAN SINGKATAN

rf	Random Forest
nb	Naive Bayes
CSV	Comma-Separated Values
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH

Analisis Sentimen	Proses mengidentifikasi opini atau ulasan dalam teks
Ulasan	Pendapat atau tanggapan pengguna
Riliv	Aplikasi kesehatan mental online dari Indonesia
Epoch	Jumlah siklus penuh pelatihan model
Optuna	Framework untuk otomatisasi pencarian hyperparameter
Visualisasi	Teknik grafis untuk melihat pola, seperti grafik, dll



INTISARI

Penelitian ini mengkaji penerapan hyperparameter tuning menggunakan Optuna untuk mengoptimalkan kinerja model algoritma Random Forest dan Naive Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Riliv. Proses klasifikasi dilakukan menggunakan teknik ekstraksi fitur TF-IDF serta penanganan ketidakseimbangan data dengan metode SMOTE. Hasil dari eksperimen ini membuktikan bahwa algoritma Random Forest memberikan performa klasifikasi lebih baik dibandingkan Naive Bayes, baik sebelum maupun sesudah dilakukan hyperparameter tuning menggunakan Optuna. Model parameter default Random Forest mampu menghasilkan nilai akurasi tertinggi sebesar 91,86% dengan nilai f1-score yang sangat baik dan seimbang disemua kelas, menunjukkan bahwa model ini sangat efektif dalam melakukan klasifikasi. Namun, setelah dilakukannya tuning hyperparameter, akurasi Random Forest justru mengalami penurunan menjadi 87,26%. Sebaliknya, algoritma Naive Bayes menunjukkan peningkatan performa setelah melakukan optimasi, dari akurasi awal sebesar 82,92% menjadi 85,22%. Meskipun demikian, performa Naive Bayes tetap berada di bawah Random Forest. Oleh karena itu, dapat disimpulkan bahwa Random Forest merupakan algoritma yang paling optimal untuk pelatihan pada penelitian ini.

Kata kunci: Klasifikasi Sentimen, Random Forest, Naive Bayes, Hyperparameter Optuna, Riliv.

ABSTRACT

This study explores of application of hyperparameter tuning using Optuna to optimize the performance of Random Forest and Naive Bayes models in classifying user sentiment reviews of the Riliv application. The classification process we carried out using TF-IDF feature extraction technique and data imbalance handling using the SMOTE method. The results of this experiment demonstrate that the Random Forest algorithm provides better classification performance than Naive Bayes, both before and after hyperparameter tuning using Optuna. The default parameter model of Random Forest achieved using Optuna. The default parameter model of Random Forest achieved the highest accuracy of 91,86% with an excellent and balanced F1-Score across all clases, indicating that this model is highly effective in classification. However, after hyperparameter tuning, the accuracy of Random Forest decreased to 87,26%. Conversely, the Naive Bayes algorithm showed an improvement in performance after optimization, from an initial accuracy of 82,92% to 85,22%. Nevertheless, the performance of Naive Bayes remained below that of Random Forest. Therefore, it can be concluded that Random Forest is the most optimal algorithm for training in this study.

Keyword: Sentiment Classification, Random Forest, Naive Bayes, Hyperparameter Optuna, Riliv.