ANALISIS PENGGUNAAN *HYPERPARAMETER TUNING* UNTUK KLASIFIKASI DATA KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *EXTREME GRADIENT BOOST*

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

CHIKA RATU MENTARI

21.11.4410

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

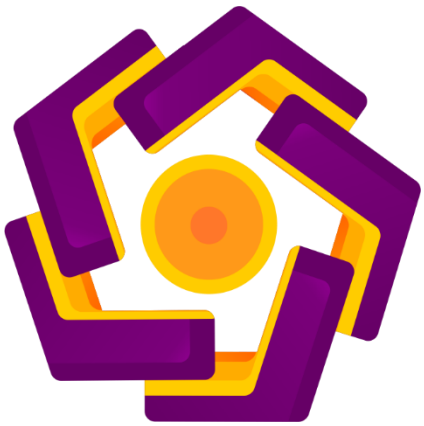
2025

ANALISIS PENGGUNAAN *HYPERPARAMETER TUNING* UNTUK KLASIFIKASI DATA KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *EXTREME GRADIENT BOOST*

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

CHIKA RATU MENTARI

21.11.4410

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2025

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS PENGGUNAAN *HYPERPARAMETER TUNNING* UNTUK KLASIFIKASI DATA KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *EXTREME GRADIENT BOOST*

yang disusun dan diajukan oleh

Chika Ratu Mentari

21.11.4410

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Februari 2025

Dosen Pembimbing,

Robert Marco, S.T., M.T., Ph.D.
NIK. 190302228

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS PENGGUNAAN *HYPERPARAMETER TUNNING* UNTUK KLASIFIKASI DATA KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *EXTREME GRADIENT BOOST*

yang disusun dan diajukan oleh

Chika Ratu Mentari

21.11.4410

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Februari 2025

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Dr. Sri Ngudi Wahyuni, ST., M.Kom. **NIK. 190302060**	
Emigawaty, M.Kom **NIK. 190302226**	
Robert Marco, S.T., M.T, Ph.D. **NIK. 190302228**	

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

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta,S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Chika Ratu Mentari
NIM : 21.11.4410

Menyatakan bahwa Skripsi dengan judul berikut:
ANALISIS PENGGUNAAN *HYPERPARAMETER TUNNING* UNTUK KLASIFIKASI DATA KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *EXTREME GRADIENT BOOST*

Dosen Pembimbing : Robert Marco, S.T., M.T., Ph.D.

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 Februari 2025

Yang Menyatakan,

Chika Ratu Mentari

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT, skripsi ini saya persembahkan kepada:

1. Allah SWT, atas limpahan rahmat, kesehatan, dan kemudahan yang diberikan sehingga skripsi ini dapat terselesaikan.
2. Kedua orang tua, yang senantiasa memberikan doa, kasih sayang, serta dukungan moral dan materi tanpa henti. Segala pencapaian ini tidak akan terwujud tanpa pengorbanan dan perjuangan kalian.
3. Dosen pembimbing dan seluruh pengajar, yang telah memberikan bimbingan, ilmu, serta arahan yang berharga dalam proses penyusunan skripsi ini. Terima kasih atas dedikasi dan kesabaran yang telah diberikan.
4. Rekan-rekan seperjuangan, yang telah menemani perjalanan akademik ini dengan dukungan, motivasi, serta kebersamaan dalam menghadapi berbagai tantangan selama studi.
5. Diriku senidiri, sebagai bentuk apresiasi atas usaha, ketekunan, dan semangat dalam menyelesaikan skripsi ini, meskipun dihadapkan pada berbagai tantangan dan rintangan.

Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi ilmu pengetahuan serta menjadi langkah awal dalam perjalanan akademik dan profesional di masa mendatang.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Fakultas Ilmu Komputer Prodi Informatika, Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, penulis menyadari bahwa keberhasilan ini tidak lepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala rasa hormat dan terima kasih, penulis ingin menyampaikan apresiasi kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom., Ph.D., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya Dhuhita, M.Kom., Selaku Ketua Program Studi Informatika Universitas Amikom Yogyakarta.
4. Bapak Robert Marco, S.T., M.T., Ph.D., selaku dosen pembimbing yang dengan sabar memberikan bimbingan, arahan, serta ilmu yang sangat berharga dalam penyusunan skripsi ini.
5. Seluruh dosen dan staf akademik Fakultas Ilmu Komputer khususnya prodi Informatika yang telah memberikan ilmu, wawasan, serta pengalaman berharga selama masa studi.
6. Teman-teman seperjuangan, yang selalu memberikan semangat, motivasi, dan kebersamaan dalam menghadapi perjalanan akademik ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Semoga skripsi ini dapat memberikan manfaat bagi pembaca serta menjadi referensi bagi penelitian selanjutnya.

Yogyakarta, 25 Februari 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 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 3
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 5
1.6 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA 6
2.1 Studi Literatur 6

2.2 Dasar Teori 10
2.2.1 Machine Learning 10
2.2.2 Klasifikasi 11
2.2.3 Algoritma Extreme Gradient Boost 12
2.2.4 *Hyperparameter* tuning 16
2.2.5 Evaluasi 18
BAB III METODE PENELITIAN 22
3.1 Alur Penelitian 22
3.3 Alat dan Bahan 32
BAB IV HASIL DAN PEMBAHASAN 34
4.1 Proses Pengumpulan Data 34
4.2 Exploratory Data Analysis (EDA) 34
4.3 Pre-processing 37
4.3 Modelling 45
4.4 Model Klasifikasi menggunakan Grid Search 46
BAB V PENUTUP 51
5.1 Kesimpulan 51
5.2 Saran 51
REFERENSI 52
LAMPIRAN 57

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian 8
Tabel 2. 2 Confution Matrix 20
Tabel 3. 1 Deskripsi Dataset 23
Tabel 3. 2 Deskripsi Statistik Data sebelum Pre-Processing 26
Tabel 3. 3 Deskripsi Statistik Data setelah Pre-Processing 29
Tabel 3. 4 Parameter algoritma yang digunakan.....31
Tabel 4. 1 Penentuan parameter terbaik.....46
Tabel 4. 2 Tabel confution matrix.....47

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian 22
Gambar 4. 1 Mental Health Dataset.....34
Gambar 4. 2 Visualisasi Data menggunakan Histogram 37
Gambar 4. 3 Sebelum Penanganan Data Duplikat.....38
Gambar 4. 4 Setelah Penanganan Data Duplikat 39
Gambar 4. 5 Sebelum Penanganan Missing Values 40
Gambar 4. 6 Sesudah Penanganan Missing Values 41
Gambar 4. 7 Sebelum proses label encoder 42
Gambar 4. 8 Sesudah proses label encoder.....42
Gambar 4. 9 Sebelum proses Normalisasi Data.....43
Gambar 4. 10 Setelah proses Normalisasi Data.....43
Gambar 4. 11 Sebelum penanganan ketidakseimbangan data 44
Gambar 4. 12 Setelah penanganan ketidakseimbangan data 44
Gambar 4. 13 Splitting data train dan data test.....45
Gambar 4. 14 Kurva ROC AUC tanpa *Hyperparameter* Tuning 48
Gambar 4. 15 Kurva ROC AUC dengan *Hyperparameter* Tuning 48
Gambar 4. 16 Confution Matrix tanpa *Hyperparameter* Tuning.....49
Gambar 4. 17 Confution Matrix dengan *Hyperparameter* Tuning.....50

DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian 10
Lampiran 2. Dokumentasi Penelitian 11

DAFTAR LAMBANG DAN SINGKATAN

TP	*True Positive*
TN	*True Negative*
FP	*False Positive*
FN	*False Negative*
XGBoost	*Extreme Gradient Boost*
LR	*Logistic Regression*
NB	*Naïve Bayes*
KNN	*K-Neighbors Nearest*
SVM	*Support Vector Machine*

DAFTAR ISTILAH

Algoritma	urutan instruksi terstruktur yang dirancang untuk memecahkan masalah tertentu secara efisien dan akurat.
Preprocessing	langkah penting dalam mempersiapkan data untuk analisis dan pembelajaran mesin, memastikan bahwa kumpulan data bersih, terstruktur, dan cocok untuk pelatihan model.
Modelling	proses dan metodologi yang digunakan untuk membuat algoritma yang dapat belajar dari data dan membuat prediksi atau keputusan.
Flowchart	alat grafis penting untuk mewakili proses dan konsep yang kompleks dengan cara yang jelas dan ringkas
Model	representasi matematis atau algoritmik dari hubungan antara data input (*features*) dan output (*target* atau *labels*).
Mayoritas	jumlah terbesar atau kelompok dominan dalam suatu populasi, komunitas, atau kumpulan data.
Minoritas	kelompok, bagian, atau jumlah yang lebih kecil dari keseluruhan dalam suatu populasi, komunitas, atau kumpulan data.
Dataset	kumpulan data yang terorganisasi dalam bentuk tertentu dan digunakan untuk berbagai keperluan analisis, pelatihan model, atau pengambilan keputusan.

INTISARI

Penelitian ini menganalisis penggunaan *hyperparameter* tuning dalam klasifikasi data kesehatan mental menggunakan algoritma *Extreme Gradient Boost* (XGBoost). Kesehatan mental merupakan aspek penting dalam kesejahteraan individu, namun data kesehatan mental sering kali tidak seimbang dan memiliki nilai yang hilang, yang dapat menghambat akurasi model klasifikasi. Untuk mengatasi masalah tersebut, penelitian ini menerapkan teknik resampling dengan *Synthetic Minority Over-sampling Technique* (SMOTE) dan imputasi data menggunakan *K-Nearest Neighbors* (KNN). Selanjutnya, dilakukan proses normalisasi dengan *MinMaxScaler* serta optimasi model menggunakan *hyperparameter tuning* dengan metode *Grid Search*. Model diuji menggunakan validasi silang *K-Fold* dan dievaluasi menggunakan confusion matrix, menghasilkan peningkatan akurasi yang signifikan dibandingkan dengan model tanpa tuning. Hasil penelitian menunjukkan bahwa kombinasi teknik *preprocessing* dan optimasi parameter dapat meningkatkan akurasi model dalam mengklasifikasikan kesehatan mental, memberikan kontribusi bagi pengembangan sistem deteksi dini yang lebih akurat.

Kata kunci: Kesehatan Mental, *Extreme Gradient Boost*, *Hyperparameter Tuning*, Klasifikasi, dan *Machine Learning*

ABSTRACT

This study analyzes the use of hyperparameter tuning in classifying mental health data using the Extreme Gradient Boost (XGBoost) algorithm. Mental health is a crucial aspect of individual well-being; however, mental health data is often imbalanced and contains missing values, which can hinder the accuracy of classification models. To address this issue, the study applies resampling techniques using the Synthetic Minority Over-sampling Technique (SMOTE) and data imputation with K-Nearest Neighbors (KNN). Furthermore, data normalization is performed using MinMaxScaler, and model optimization is conducted through hyperparameter tuning with the Grid Search method. The model is tested using K-Fold cross-validation and evaluated with a confusion matrix, resulting in a significant accuracy improvement compared to models without tuning. The findings indicate that the combination of preprocessing techniques and parameter optimization enhances model accuracy in classifying mental health conditions, contributing to the development of a more accurate early detection system.

Keyword: *Mental Health, Extreme Gradient Boost, Hyperparameter Tuning, Classification, Machine Learning.*