

**PERBANDINGAN KINERJA EXTREME GRADIENT BOOST
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
TINGKAT OBESITAS BERDASARKAN POLA HIDUP
SEHARI-HARI**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S-1 Informatika



disusun oleh

MAGHFIROH IZZAH MAULIYANA

21.11.3862

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**PERBANDINGAN KINERJA EXTREME GRADIENT BOOST
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
TINGKAT OBESITAS BERDASARKAN POLA HIDUP
SEHARI-HARI**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S-I Informatika



disusun oleh

MAGHFIROH IZZAH MAULIYANA

21.11.3862

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**PERBANDINGAN KINERJA EXTREME GRADIENT BOOST
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
TINGKAT OBESITAS BERDASARKAN POLA HIDUP
SEHARI-HARI**

yang disusun dan diajukan oleh

Maghfiroh Izzah Mauliyana

21.11.3862

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Agustus 2025

Dosen Pembimbing,



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

NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

**PERBANDINGAN KINERJA EXTREME GRADIENT BOOST
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
TINGKAT OBESITAS BERDASARKAN POLA HIDUP**

SEHARI-HARI

yang disusun dan diajukan oleh

Maghfiroh Izzah Mauliyana

21.11.3862

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Nuri Cahyono, M.Kom.
NIK. 190302278

Haryoko, S.Kom., M.Cs.
NIK. 190302286

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

Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Maghfiroh Izzah Mauliyana
NIM : 21.11.3862

Menyatakan bahwa Skripsi dengan judul berikut:

**PERBANDINGAN KINERJA EXTREME GRADIENT BOOST DAN
LOGISTIC REGRESSION DALAM MEMPREDIKSI TINGKAT
OBESITAS BERDASARKAN POLA HIDUP SEHARI-HARI**
Dosen Pembimbing : Windha Mega Pradnya Dhuhita, 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, 19 Agustus 2025

Yang Menyatakan,



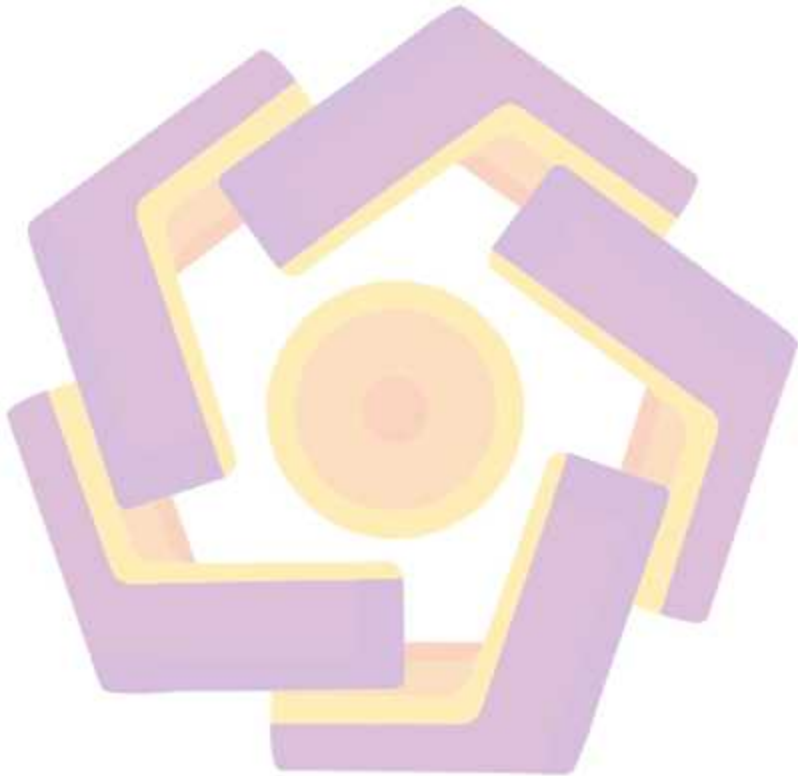
Maghfiroh Izzah Mauliyana

HALAMAN PERSEMBAHAN

Segala puji bagi Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini. Selama penyusunan laporan ini, penulis memperoleh berbagai bentuk dukungan, bimbingan, dan bantuan dari banyak pihak. Dengan demikian, skripsi ini penulis dedikasikan kepada:

1. Kepada cinta pertama dan pintu surga penulis, ayah dan ibu. Penulis ucapkan terima kasih atas segala doa dan dukungan yang tidak pernah putus. Memberikan cinta, kasih sayang, doa, dan pengorbanan yang mengiringi setiap langkah untuk menyelesaikan pendidikan ini. Terima kasih telah mengantarkan penulis sampai di titik ini. Terima kasih sudah berjuang untuk penulis, membesarkan dan mendidik penulis sampai mendapatkan gelar sarjana.
2. Kepada adik laki-laki penulis, yang selalu membuat penulis termotivasi untuk bisa terus belajar menjadi sosok kakak yang dapat memberikan pengaruh positif, baik dalam bidang akademik maupun non-akademik, serta berusaha menjadi panutan di masa yang akan datang kelak.
3. Kepada sahabat seperjuangan dan seperantauan yaitu Salma, Saroh, Christa dan Pur. Terima kasih sudah menemani penulis selama masa perkuliahan. Terima kasih sudah menjadi teman, sahabat, dan saudara yang selalu menguatkan penulis. Mengambil banyak peran penting dibalik layar, kebersamaan dalam perjuangan dan tidak pernah mengeluh saat direpotkan.
4. Kepada rekan seperjuangan di angkatan 2021 program studi S1-Informatika, Universitas Amikom Yogyakarta. Khususnya kelas 21IF01 dan pasukan Umi Kultsum, yang selalu memberikan dukungan dan motivasi sejak semester satu perkuliahan hingga akhirnya dapat lulus bersama-sama. Meskipun setelah ini akan menjalani kehidupan masing-masing yang berbeda, kesibukan yang berbeda-beda dan mungkin berada di kota atau negara yang berbeda, semoga pertemanan ini selalu terjaga selamanya.

5. Kepada semua pihak baik yang mendukung secara langsung maupun tidak langsung, yang telah membantu penulis dalam menyelesaikan skripsi ini.



KATA PENGANTAR

Puji dan syukur kita panjatkan kepada Allah Subhanahu wa Ta'ala, Tuhan yang Maha Esa, yang telah melimpahkan rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi yang berjudul "Perbandingan Kinerja Extreme Gradient Boost dan Logistic Regression dalam Memprediksi Tingkat Obesitas Berdasarkan Pola Hidup Sehari-Hari" ini disusun sebagai salah satu syarat untuk menyelesaikan program studi S1-Informatika di Universitas Amikom Yogyakarta.

Penulis menyadari bahwa penyelesaian skripsi ini tidak lepas dari bantuan, dukungan, serta bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Allah Subhanahu wa Ta'ala, yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Prof. Dr .M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta.
3. Ibu Prof. Dr. Kusriani, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
4. Ibu Eli Pujiastuti, M.Kom. selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
5. Ibu Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom. selaku dosen pembimbing, yang dengan sabar memberikan bimbingan, arahan, dan motivasi yang sangat berharga selama proses penyusunan skripsi ini.
6. Bapak Nuri Cahyono, M.Kom., dan Bapak Haryoko, S.Kom., M.Cs., selaku dosen penguji yang telah memberikan saran dan masukan atas penyelesaian skripsi ini.
7. Kepada seluruh Dosen, Staff, dan Karyawan Universitas Amikom Yogyakarta, yang telah memberikan ilmu pengetahuan yang bermanfaat selama penulis menjalani perkuliahan serta membantu kelancaran administrasi yang mendukung penyelesaian skripsi ini.

8. Keluarga, sahabat dan teman-teman yang senantiasa memberikan dukungan, semangat dan doa.
9. Serta kepada semua pihak yang telah memberikan bantuan dalam penyusunan skripsi ini, yang tidak dapat penulis sebutkan satu per satu.

Penulis berharap skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, terutama di bidang Machine Learning, serta bagi pihak-pihak yang membutuhkan. Meskipun demikian, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan untuk perbaikan di masa depan.

Yogyakarta, 19 Agustus 2025

Penulis

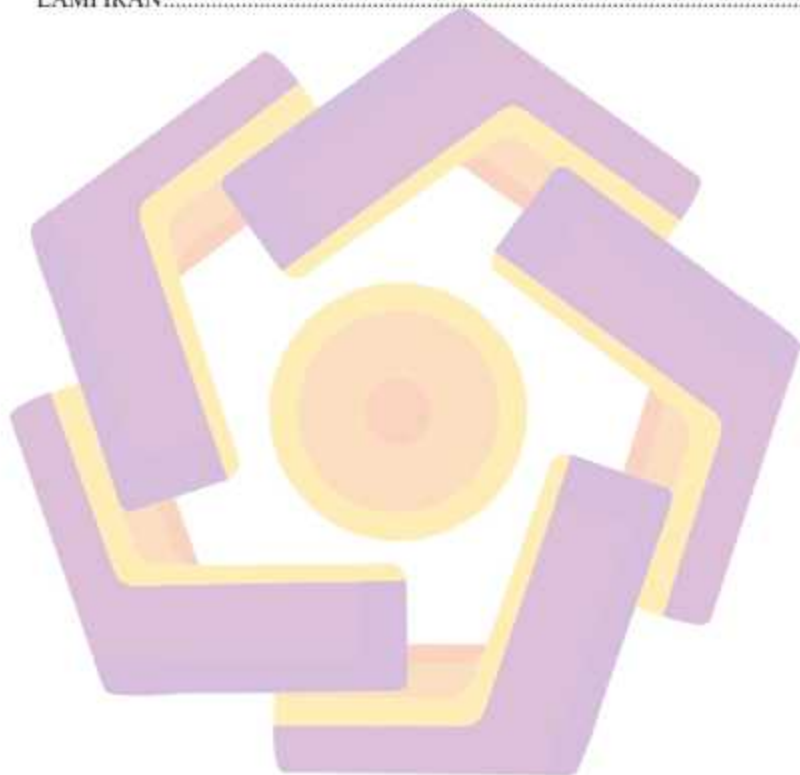


DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
INTISARI	xv
<i>ABSTRACT</i>	xvi
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 Obesitas.....	12
2.2.2 Pola Hidup	13

2.2.3 Data Mining	14
2.2.4 Machine Learning	14
2.2.5 Data Preprocessing	15
2.2.6 Extreme Gradient Boosting (XGBoost)	17
2.2.7 Logistik Regression	18
2.2.8 Hyperparameter Tuning	18
2.2.9 Evaluasi	19
2.2.10 Deployment	21
BAB III METODE PENELITIAN	22
3.1 Objek Penelitian	22
3.2 Alur Penelitian	23
3.2.1 Pengambilan data	23
3.2.2 Data Preprocessing	25
3.2.3 Modelling	28
3.2.4 Evaluasi	30
3.2.5 Deployment	30
3.3 Alat dan Bahan	32
3.3.1 Data Penelitian	32
3.3.2 Alat/instrumen	32
BAB IV HASIL DAN PEMBAHASAN	33
4.1 Pengambilan data	33
4.3 Preprocessing	34
4.4 Model	39
4.5 Evaluasi	45
4.6 Deployment	55

BAB V PENUTUP	58
5.1 Kesimpulan	58
5.2 Saran	58
REFERENSI	60
LAMPIRAN.....	65



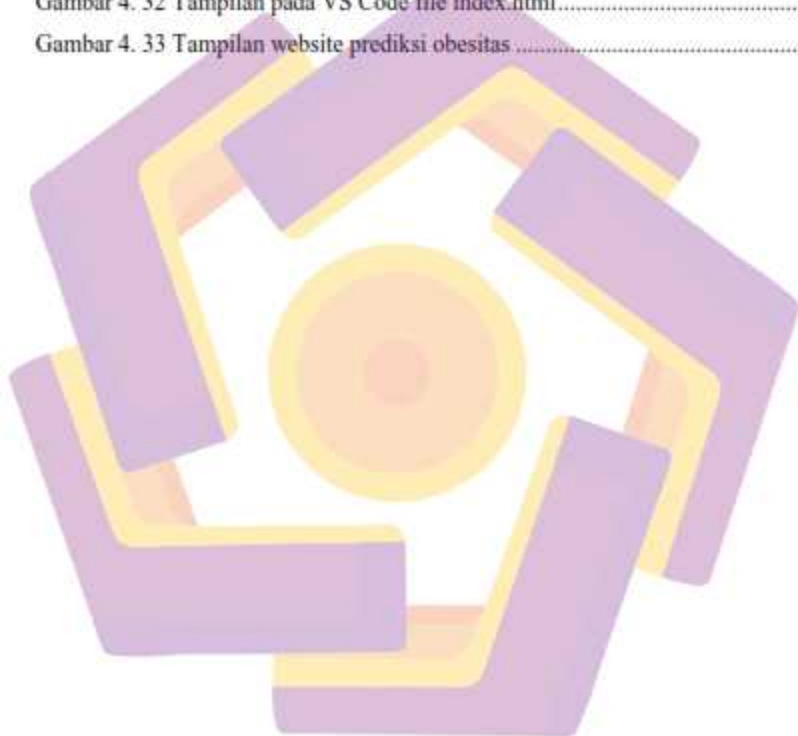
DAFTAR TABEL

Tabel 2. 1 Keaslian penelitian.....	8
Tabel 2. 2 Confussion Matrix	19
Tabel 2. 3 Confussion Matrix Multi-Class	20
Tabel 3. 1 Deskripsi fitur	24
Tabel 3. 2 Hyperparameter Tuning XGBoost.....	29
Tabel 3. 3 Hyperparameter Tuning Logistic Regression.....	30
Tabel 3. 4 Spesifikasi Hardware	32
Tabel 3. 5 Spesifikasi Software.....	32
Tabel 4. 1 Sebelum standarisasi.....	38
Tabel 4. 2 Setelah standarisasi.....	38
Tabel 4. 3 Parameter baseline Logistic Regression	40
Tabel 4. 4 Parameter baseline XGBoost.....	41
Tabel 4. 5 Hyperparameter Tuning Logistic Regression	42
Tabel 4. 6 Hyperparameter Tuning XGBoost.....	45

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	23
Gambar 3. 2 Alur sistem website	31
Gambar 3. 3 Wireframe website	31
Gambar 4. 1 5 baris tearatas dataframe	33
Gambar 4. 2 Informasi Dataframe	33
Gambar 4. 3 Penambahan fitur BMI	34
Gambar 4. 4 Fitur BMI	34
Gambar 4. 5 Manual mapping fitur target	35
Gambar 4. 6 Dataframe setelah manual mapping	35
Gambar 4. 7 Label encoding	35
Gambar 4. 8 Dataframe setelah label encoding	36
Gambar 4. 9 One hot encoding	36
Gambar 4. 10 Hasil one hot encoding	37
Gambar 4. 11 Split dataset	37
Gambar 4. 12 Hasil distribusi split dataset	37
Gambar 4. 13 Standarisasi	37
Gambar 4. 14 SMOTE	39
Gambar 4. 15 Distribusi target sebelum dan sesudah SMOTE	39
Gambar 4. 16 Model Logistic Regression	40
Gambar 4. 17 Model XGBoost	41
Gambar 4. 18 Model Logistic Regression dengan Hyperparameter Tuning	42
Gambar 4. 19 Model XGBoost dengan Hyperparameter Tuning	44
Gambar 4. 20 Classification report Logistic Regression default	45
Gambar 4. 21 Confussion Matrix Logistic Regression default	46
Gambar 4. 22 Classification report XGBoost default	47
Gambar 4. 23 Confussion Matrix XGBoost default	48
Gambar 4. 24 Classification report Logistic Regression dengan Hyperparameter Tuning	49
Gambar 4. 25 Confussion matrix Logistic Regression dengan Hyperparameter Tuning	50

Gambar 4. 26 Classification report XGBoost dengan Hyperparameter Tuning	51
Gambar 4. 27 Confussion matrix XGBoost dengan Hyperparameter Tuning	52
Gambar 4. 28 Grafik perbandingan kinerja model	53
Gambar 4. 29 Perbandingan waktu pelatihan	54
Gambar 4. 30 Perbandingan penggunaan memori	54
Gambar 4. 31 Tampilan pada VS Code file app.py	55
Gambar 4. 32 Tampilan pada VS Code file index.html	56
Gambar 4. 33 Tampilan website prediksi obesitas	56



INTISARI

Penelitian ini menganalisis dan membandingkan kinerja dua algoritma *machine learning*, yaitu Extreme Gradient Boosting (XGBoost) dan Logistic Regression, dalam memprediksi tingkat obesitas berdasarkan pola hidup sehari-hari. Obesitas merupakan masalah kesehatan global yang terus meningkat dan memiliki dampak besar terhadap kesehatan individu serta sistem kesehatan masyarakat. Deteksi dini melalui analisis pola hidup menjadi hal yang sangat penting dalam pencegahan. Penelitian ini menggunakan dataset "Obesity Prediction Dataset" yang mencakup faktor-faktor seperti kebiasaan makan, aktivitas fisik, dan demografis. Kedua model dievaluasi dengan parameter default dan melalui proses *hyperparameter tuning*. Hasil penelitian menunjukkan bahwa XGBoost lebih unggul, dengan akurasi 99,06% setelah *tuning*, dibandingkan Logistic Regression yang mencapai akurasi 95,28%. Meskipun XGBoost memberikan akurasi yang lebih tinggi, model ini memiliki biaya komputasi yang lebih tinggi, dengan waktu pelatihan dan penggunaan memori yang lebih besar. Sementara itu, Logistic Regression lebih efisien namun memberikan hasil prediksi yang lebih rendah. Penelitian ini menyimpulkan bahwa XGBoost merupakan algoritma yang paling efektif untuk memprediksi obesitas secara akurat, sementara Logistic Regression menawarkan solusi yang lebih praktis untuk lingkungan dengan sumber daya terbatas.

Kata kunci: Obesitas, XGBoost, Logistic Regression, Hyperparameter Tuning, Machine Learning

ABSTRACT

This study analyzes and compares the performance of two machine learning algorithms, Extreme Gradient Boosting (XGBoost) and Logistic Regression, in predicting obesity levels based on daily lifestyle patterns. Obesity has become a growing global health issue with significant impacts on individual health and healthcare systems. Early detection through lifestyle pattern analysis is crucial for prevention. The study utilizes the "Obesity Prediction Dataset" which includes factors such as eating habits, physical activity, and demographic. Both models were evaluated with default parameters and through hyperparameter tuning. The results indicate that XGBoost outperforms Logistic Regression, achieving an accuracy of 99.06% after tuning, compared to 95.28% for Logistic Regression. Although XGBoost provides higher accuracy, it comes with a higher computational cost, requiring more training time and memory. In contrast, Logistic Regression is more efficient but offers lower predictive performance. This research concludes that XGBoost is the most effective algorithm for accurate obesity prediction, while Logistic Regression provides a more practical solution for resource-limited environments.

Keyword: *Obesity, XGBoost, Logistic Regression, Hyperparameter Tuning, Machine Learning*