

**MENGANALISIS KETERGANTUNGAN ANTROPOMETRIK
DALAM KLASIFIKASI OBESITAS MENGGUNAKAN ABLASI
FITUR**

SKRIPSI NON REGULER JALUR SCIENTIST

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



Disusun oleh :

ESTI DWI PUSPITASARI

22.11.4988

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**MENGANALISIS KETERGANTUNGAN ANTROPOMETRIK
DALAM KLASIFIKASI OBESITAS MENGGUNAKAN ABLASI
FITUR**

SKRIPSI NON REGULER JALUR SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



Disusun oleh :

ESTI DWI PUSPITASARI

22.11.4988

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2020

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**MENGANALISIS KETERGANTUNGAN ANTROPOMETRIK DALAM
KLASIFIKASI OBESITAS MENGGUNAKAN ABLASI FITUR**

yang disusun dan diajukan oleh

Esti Dwi Puspitasari

22.11.4988

telah disetujui oleh Dosen Pembimbing
pada tanggal 11 Mei 2026

Dosen Pembimbing,



Bayu Setiaji, M.Kom

NIK. 190302216

HALAMAN PENGESAHAN

JALUR NON-REGULER

**MENGANALISIS KETERGANTUNGAN ANTROPOMETRIK DALAM
KLASIFIKASI OBESITAS MENGGUNAKAN ABLASI FITUR**

yang disusun dan diajukan oleh

Esti Dwi Puspitasari
22.11.4988

Telah dipertahankan di depan Dewan Penguji
pada tanggal 11 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Arifivanto Hadinegoro, S.Kom., M.T.
NIK. 190302289



Nuri Cahyono, M.Kom.
NIK. 190302278



Bayu Setiaji, M.Kom.
NIK. 190302216



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 11 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Esti Dwi Puspitasari

NIM : 22.11.4988

Menyatakan bahwa Laporan dengan judul berikut:

Menganalisis Keteranganfungan Antropometrik Dalam Klasifikasi Obesitas Menggunakan Ablasi Fitur

Dosen Pembimbing : Bapak Bayu Setiaji, 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 kegiatan 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, 11 Mei 2026



UNIVERSITAS AMIKOM
YOGYAKARTA
METRA
TAMBAH
008E00X40247283

Esti Dwi Puspitasari

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT atas segala rahmat dan petunjuk-Nya, karya ini penulis persembahkan kepada kedua orang tua dan keluarga tercinta yang senantiasa memberikan doa, dukungan, dan kasih sayang yang tulus tanpa henti. Sahabat dan teman-teman, yang telah memberikan dukungan, kebersamaan, serta semangat selama penyusunan laporan ini. Semua pihak yang telah membantu, baik secara langsung maupun tidak langsung, dalam penyelesaian laporan ini.



KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa atas rahmat, karunia serta petunjuk-Nya sehingga penulis dapat menyelesaikan laporan yang berjudul “Menganalisis Ketergantungan Antropometrik Dalam Klasifikasi Obesitas Menggunakan Ablasi Fitur”. Laporan ini menjadi bagian dari salah satu syarat bagi penulis untuk memperoleh gelar Sarjana Komputer di Universitas Amikom Yogyakarta.

Penulis menyadari bahwa dalam proses penyusunan dan penyelesaian karya ini tidak terlepas dari dukungan, bantuan dan doa dari berbagai pihak. Oleh karena itu dengan segala kerendahan hati, penulis mengucapkan terima kasih kepada:

1. Bapak Bayu Setiaji, M. Kom, selaku dosen pembimbing yang telah dengan sabar membimbing, memberikan arahan, dan masukan berharga, serta meluangkan waktu selama proses penyusunan jurnal hingga penyusunan laporan ini.
2. Prof. Dr. M. Suyanto., M.M, selaku rektor Universitas Amikom Yogyakarta.
3. Prof. Dr. Kusnini, M.Kom, selaku Dekan Fakultas Ilmu Komputer
4. Ibu Eli Pujiastuti, M.Kom, selaku ketua Program Studi Informatika.
5. Seluruh dosen dan staff karyawan Universitas Amikom Yogyakarta, atas ilmu, fasilitas dan pelayanan terbaik yang diberikan selama masa perkuliahan penulis.
6. Bapak Dr. Cipto Suncoko, M.Pd dan Ibu Hamalia Wado, selaku kedua orang tua penulis yang senantiasa memberikan doa, kasih sayang, semangat, dan dukungan baik moral maupun material.
7. Kepada Cilia, Juninda, dan Aiko, selaku saudara kandung penulis yang telah memberikan semangat, motivasi, dukungan serta kepercayaan yang sangat berarti dalam mendukung kelancaran studi penulis.

8. Teman-teman seperjuangan yang penulis sayangi Deka, Fifi, Nisa, serta Aulia selaku sahabat dekat penulis, dan teman-teman yang tidak bisa penulis sebutkan satu persatu. Terima kasih atas semangat, dukungan, hiburan dan kebersamaan yang telah diberikan selama ini.

Penulis menyadari bahwa laporan ini masih jauh dari kata sempurna, mengingat adanya keterbatasan yang dimiliki penulis dalam proses penyusunannya. Oleh karena itu, dengan segala kerendahan hati, penulis mengharapkan kritik dan saran yang membangun guna penyempurnaan di masa yang akan datang.

Akhir kata, penulis berharap laporan ini dapat memberikan manfaat bagi para pembaca serta berkontribusi dalam pengembangan ilmu pengetahuan, khususnya di bidang Machine Learning.

Yogyakarta, 12 Mei 2026

Penulis



DAFTAR ISI

Halaman Judul.....	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN KARYA.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI.....	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN	1
1.1. Gambaran Umum	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	4
BAB II TEORI DAN METODE.....	5
2.1. Teori	5
2.1.1. Antropometrik	5
2.1.2. Machine Learning.....	5
2.1.3. Klasifikasi Obesitas	6
2.1.4. Ablasi Fitur.....	6
2.1.5. Logistic Regression	7
2.1.6. Random Forest	7
2.1.7. Data Preprocessing	8
2.1.8. Confusion Matrix	8
2.1.9. Evaluasi Matriks.....	10

2.2.	Analisis	10
2.2.1.	Mengidentifikasi Masalah	10
2.2.2.	Mencari Jurnal yang Relevan	11
2.2.3.	Dataset	11
2.2.4.	Exploratory Data Analysis (EDA)	12
2.2.5.	Data Cleaning dan Quality Control	15
2.2.6.	Pemisahan Target dan Fitur	15
2.2.7.	Pembagian Data	15
2.2.8.	Data Preprocessing	15
2.2.9.	Feature Configuration	16
2.2.10.	Model yang Digunakan	17
2.2.11.	Training dan Prediction	18
2.2.12.	Evaluasi Model	19
2.2.13.	Error Analysis	19
2.3.	Alur Pengembangan	19
BAB III HASIL DAN PEMBAHASAN		23
3.1.	Hasil	23
3.2.	Evaluasi	45
BAB IV KESIMPULAN		50
4.1.	Kesimpulan	50
4.2.	Saran	51
REFERENSI		52
CURICULUM VITAE		56
LAMPIRAN DAN BUKTI PENDUKUNG		58

DAFTAR TABEL

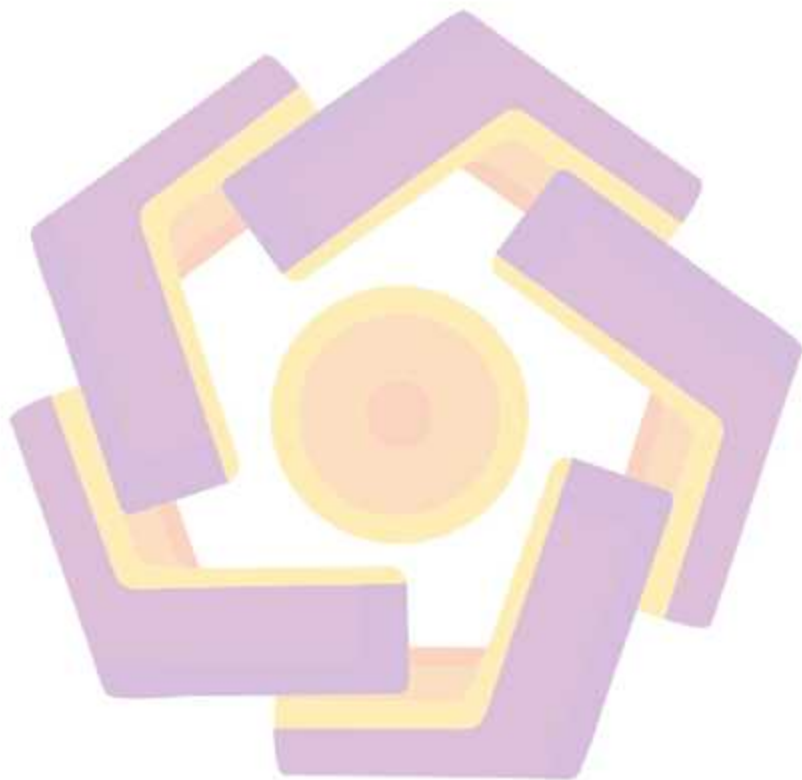
Tabel 2. 1 Confusion Matrix.....	9
Tabel 2. 2 Fitur Dataset	12
Tabel 2. 3 Tipe data Dataset	14
Tabel 3. 1 Hasil Performa LR pada Baseline Tiap Kelas	25
Tabel 3. 2 LR Baseline Classification Report	26
Tabel 3. 3 Hasil Performa LR pada Ablation Tanpa Weight dan Height.....	27
Tabel 3. 4 LR Ablation Classification Report	28
Tabel 3. 5 Hasil Performa RF pada Baseline Tiap Kelas	30
Tabel 3. 6 RF Baseline Classification Repprot.....	31
Tabel 3. 7 Hasil Performa LR pada Ablation Tanpa Weight dan Height.....	32
Tabel 3. 8 RF Ablation Classification Report.....	32
Tabel 3. 9 Hasil Performa LR pada Ablation Tanpa Weight Tiap Kelas	35
Tabel 3. 10 LR Tanpa Weight Classification Report	36
Tabel 3. 11 Hasil Performa RF Tanpa Weight Tiap Kelas	38
Tabel 3. 12 RF Tanpa Weight Classification Report	39
Tabel 3. 13 Hasil Performa LR Tanpa Height Tiap Kelas	41
Tabel 3. 14 LR Tanpa Height Classification Report.....	41
Tabel 3. 15 Hasil Performa RF Tanpa Height Tiap Kelas	44
Tabel 3. 16 RF Tanpa Height Classification Report.....	45

DAFTAR GAMBAR

Gambar 2. 1 Dataset Nobeyesdad	12
Gambar 2. 2 Alur Penelitian	22
Gambar 3. 1 Hasil Klasifikasi LR pada Baseline (semua fitur) dan Ablation (tanpa Weight dan Height).....	24
Gambar 3. 2 Grafik Hasil Klasifikasi LR.....	24
Gambar 3. 3 Confusion Matrix LR dengan semua fitur.....	25
Gambar 3. 4 Confusion Matrix LR tanpa Wight dan Height	27
Gambar 3. 5 Hasil Klasifikasi RF pada skenario I Baseline (semua fitur) dan Ablation (tanpa Weight dan Height).....	29
Gambar 3. 6 Grafik Hasil Klasifikasi RF	29
Gambar 3. 7 Confusion Matrix RF semua fitur.....	30
Gambar 3. 8 Confusion Matrix RF tanpa Weight dan Height	32
Gambar 3. 9 Hasil Klasifikasi LR pada Skenario I tanpa fitur Weight.....	34
Gambar 3. 10 Grafik Perbandingan LR tanpa fitur Weight	34
Gambar 3. 11 Confusion Matrix tanpa fitur Weight	35
Gambar 3. 12 Hasil Klasifikasi RF pada Skenario II tanpa Weight.....	36
Gambar 3. 13 Grafik Perbandingan RF tanpa Weight.....	37
Gambar 3. 14 Confusion Matrix tanpa Weight	37
Gambar 3. 15 Hasil Klasifikasi LR pada Skenario III tanpa fitur Height.....	39
Gambar 3. 16 Perbandingan Hasil Klasifikasi LR Skenario III	40
Gambar 3. 17 Confusion Matrix Tanpa Fitur Height	41
Gambar 3. 18 Hasil Klasifikasi RF pada Skenario III tanpa Fitur Height	42
Gambar 3. 19 Hasil Perbandingan Klasifikasi RF tanpa Fitur Height.....	43
Gambar 3. 20 Confusion Matrix RF tanpa Fitur Height	44

DAFTAR LAMPIRAN

Lampiran 1 Letter of Acceptance (LOA)	58
Lampiran 2 Lembar Review	59



DAFTAR LAMBANG DAN SINGKATAN

x_i	Nilai fitur ke- i sebelum encoding
c_k	Kelas ke- k dari suatu fitur
k	Jumlah kelas
K	Jumlah total kelas
$P(y = k x)$	Probabilitas data dengan fitur x termasuk kelas ke- k
x	Vektor fitur (input)
y	Label atau kelas target
e	Konstanta eksponensial (≈ 2.718)
j	Jumlah untuk penjumlahan seluruh kelas
FAVC	<i>Frequent Consumption of High Caloric Food</i>
FCVC	<i>Frequency of Consumption of Vegetables</i>
NCP	<i>Number of Main Meals</i>
CAEC	<i>Consumption of Food Between Meals</i>
SMOKE	<i>Smoking Habit</i>
CH2O	<i>Consumption of Water (H₂O)</i>
SCC	<i>Monitoring of Calorie Consumption</i>
FAF	<i>Frequency of Physical Activity</i>
TUE	<i>Time Using Electronic Devices</i>
CALC	<i>Consumption of Alcohol</i>
MTRANS	<i>Mode of Transportation</i>
LR	<i>Logistic Regression</i>
RF	<i>Random Forest</i>
TP	<i>True Positive</i>
TN	<i>True Negative</i>
FP	<i>False Positive</i>
FN	<i>False Negative</i>
BMI	<i>Body Mass Index</i>

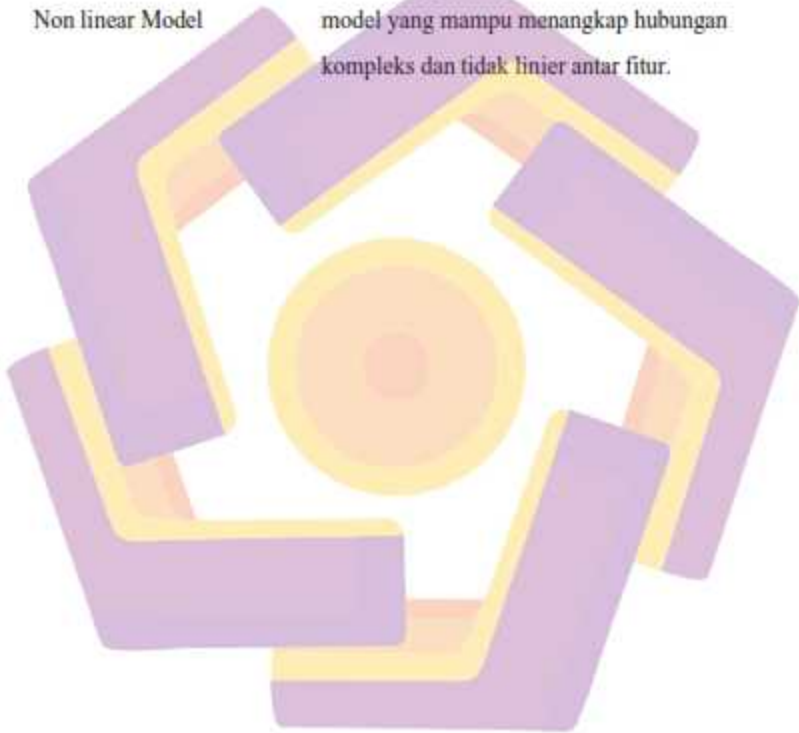
DAFTAR ISTILAH

Ablation Fitur	metode analisis dalam machine learning yang dilakukan dengan menghilangkan satu atau beberapa fitur untuk mengevaluasi pengaruhnya terhadap performa model.
Anthropometric	berkaitan dengan pengukuran fisik tubuh manusia, seperti tinggi badan, berat badan, dan indeks massa tubuh.
Klasifikasi	proses pengelompokan data ke dalam kategori tertentu berdasarkan pola yang dipelajari oleh model.
Feature Dependency	tingkat ketergantungan model terhadap fitur tertentu dalam menghasilkan prediksi.
Obesitas	kondisi medis yang ditandai dengan kelebihan lemak tubuh yang dapat meningkatkan risiko berbagai penyakit.
Feature	variabel atau atribut yang digunakan sebagai input dalam model machine learning.
Logistic Regression	algoritma klasifikasi berbasis model linear yang digunakan untuk memprediksi kemungkinan suatu kelas.
Random Forest	Algoritma ensemble berbasis decision tree yang menggabungkan banyak pohon dan menggabungkan hasil prediksi melalui pemungutan suara.
Machine Learning	bagian kecerdasan buatan yang memungkinkan sistem belajar dari data dan membuat prediksi tanpa diprogram secara langsung.
Data Preprocessing	tahapan pengolahan data sebelum digunakan dalam model, seperti pembersihan data, normalisasi, dan transformasi.



Pipeline	serangkaian proses terstruktur yang menggabungkan preprocessing dan pemodelan dalam satu alur kerja.
StandartScaler	metode normalisasi data yang mengubah fitur agar memiliki rata-rata 0 dan standar deviasi 1.
Confusion Matrix	matriks evaluasi yang digunakan untuk membandingkan hasil prediksi model dengan nilai aktual.
One-Hot Encoding	teknik mengubah data kategorikal menjadi numerik dalam bentuk vektor biner.
Feature Configuration	pengaturan kombinasi fitur yang digunakan dalam eksperimen model, termasuk skenario baseline dan ablation.
Error Analysis	proses analisis kesalahan prediksi model untuk memahami pola kesalahan yang terjadi.
Balanced Accuracy	rata-rata akurasi pada setiap kelas, digunakan untuk menangani data yang tidak seimbang.
Macro F1-score	rata-rata nilai F1-score dari setiap kelas tanpa mempertimbangkan distribusi jumlah data.
Accuracy	persentase jumlah prediksi yang benar dibandingkan dengan total data.
Precision	tingkat ketepatan model dalam memprediksi kelas positif.
Recall	kemampuan model dalam menemukan seluruh data yang benar-benar positif.
F1-score	nilai antara precision dan recall yang digunakan untuk mengukur keseimbangan keduanya.
True Positive	data positif yang diprediksi dengan benar sebagai Positif.

True Negative	data negatif yang diprediksi dengan benar sebagai negatif.
False Positive	data negatif yang salah diprediksi sebagai positif.
False Negative	data positif yang salah diprediksi sebagai negatif.
Linear Model	model yang mengasumsikan hubungan linier antara fitur dan target.
Non linear Model	model yang mampu menangkap hubungan kompleks dan tidak linier antar fitur.



INTISARI

Studi ini menganalisis ketergantungan model klasifikasi obesitas pada fitur antropometrik, khususnya pada fitur Berat dan Tinggi, yang digunakan secara langsung untuk menghitung indeks massa tubuh (BMI). Studi ini melakukan ablasi dan analisis kesalahan menggunakan model Regresi Logistik dan Random Forest dalam tiga skenario, yaitu skenario 1 baseline semua fitur vs ablasi tanpa fitur Berat dan Tinggi, skenario 2 ablasi parsial tanpa fitur Berat, dan skenario 3 ablasi parsial tanpa fitur Tinggi. Hasil evaluasi menggunakan Akurasi Seimbang, skor F1 Makro, *confusion matrix*, dan metrik per kelas. Model Random Forest pada skenario 1 mencapai Akurasi Seimbang 0,95 dan skor F1 Makro 0,94, sedangkan Regresi Logistik mencapai 0,93 dan 0,92. Selanjutnya, pada skenario 2, hasil evaluasi menjadi 0,88 dan F1 Makro 0,87 untuk Random Forest, kemudian hasil evaluasi menjadi 0,85 dan F1 Makro 0,84 untuk Regresi Logistik. Studi ablasi parsial menunjukkan bahwa penghapusan fitur Berat memiliki dampak yang lebih besar pada penurunan kinerja dibandingkan dengan penghapusan fitur Tinggi, yang mengindikasikan bahwa model lebih bergantung pada data berat badan. Analisis kesalahan menunjukkan peningkatan kebingungan antara kelas obesitas yang berdekatan, khususnya antara kategori kelebihan berat badan dan obesitas tipe I. Temuan ini menunjukkan bahwa sebagian besar kemampuan prediksi dalam klasifikasi obesitas masih dipengaruhi oleh fitur yang terkait dengan BMI, oleh karena itu analisis ketergantungan fitur sangat penting dalam mengembangkan model klasifikasi obesitas yang lebih kuat.

Kata kunci: Studi Ablasi, Ketergantungan Fitur, Klasifikasi Obesitas, Random Forest, Regresi Logistik.

ABSTRACT

This study analyzes the dependence of obesity classification models on anthropometric features, especially on Weight and Height features, which are used directly to calculate body mass index (BMI). This study conducted ablation and error analysis using Logistic Regression and Random Forest models in three scenarios, namely scenario 1 baseline all features vs ablation without Weight and Height features, scenario 2 partial ablation without Weight features, and scenario 3 partial ablation without Height features. The evaluation results used Balanced Accuracy, Macro F1-score, confusion matrix, and per-class metrics. The Random Forest model in scenario 1 achieved a Balanced Accuracy of 0.95 and a Macro F1-score of 0.94, while Logistic Regression achieved 0.93 and 0.92. Furthermore, in scenario 2, the evaluation results became 0.88 and Macro F1 0.87 for Random Forest, then the evaluation results were 0.85 and Macro F1 0.84 for Logistic Regression. Partial ablation studies show that removing the Weight feature has a greater impact on performance degradation compared to removing the Height feature, indicating that the model relies more on weight data. Error analysis shows increased confusion between adjacent obesity classes, particularly between the overweight and type I obesity categories. These findings indicate that a significant portion of the predictive ability in obesity classification is still influenced by features related to BMI, therefore feature dependency analysis is crucial in developing more robust obesity classification models.

Keyword: Ablation Study, Feature Dependency, Obesity Classification, Random Forest, Logistic Regression