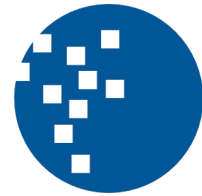


**RANCANG BANGUN SISTEM KLASIFIKASI TINGKAT OBESITAS
DENGAN ALGORITMA *RANDOM FOREST CLASSIFIER***



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Tesalonika Abigail
00000044503

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024**

**RANCANG BANGUN SISTEM KLASIFIKASI TINGKAT OBESITAS
DENGAN ALGORITMA *RANDOM FOREST CLASSIFIER***



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Tesalonika Abigail

00000044503

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Tesalonika Abigail

Nomor Induk Mahasiswa : 00000044503

Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Skripsi saya yang berjudul:
**Rancang Bangun Sistem Klasifikasi Tingkat Obesitas dengan Algoritma
*Random Forest Classifier***

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 07 Mei 2024



(Tesalonika Abigail)

HALAMAN PENGESAHAN

Skripsi dengan judul

RANCANG BANGUN SISTEM KLASIFIKASI TINGKAT OBESITAS DENGAN ALGORITMA *RANDOM FOREST CLASSIFIER*

oleh

Nama : Tesalonika Abigail
NIM : 00000044503
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 28 Mei 2024

Pukul 10.00 s/s 12.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

Penguji

(Dr. Ir. Winarno, M.Kom.)

NIDN: 0330106002

(Sy Yuliani Yakub, S.Kom., M.T. PhD)

NIDN: 0411037904

Pembimbing

(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

PJS Ketua Program Studi Informatika,

(Yaman Khaeruzzaman, M.Sc.)

NIDN: 0413057104

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Tesalonika Abigail
NIM : 00000044503
Program Studi : Informatika
Jenjang : S1
Jenis Karya : Skripsi

Menyatakan dengan sesungguhnya bahwa saya bersedia:

- ✓ Memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.

Saya tidak bersedia, dikarenakan:

- Dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 07 Mei 2024

Yang menyatakan



Tesalonika Abigail

U M M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

** Jika tidak bisa membuktikan LoA jurnal/HKI selama 6 bulan kedepan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan/ Motto

"You will definitely enjoy what you've worked hard for-you'll be happy; and things will go well for you."

Psalm 128:2



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Rancang Bangun Sistem Klasifikasi Tingkat Obesitas dengan Algoritma *Random Forest Classifier* dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

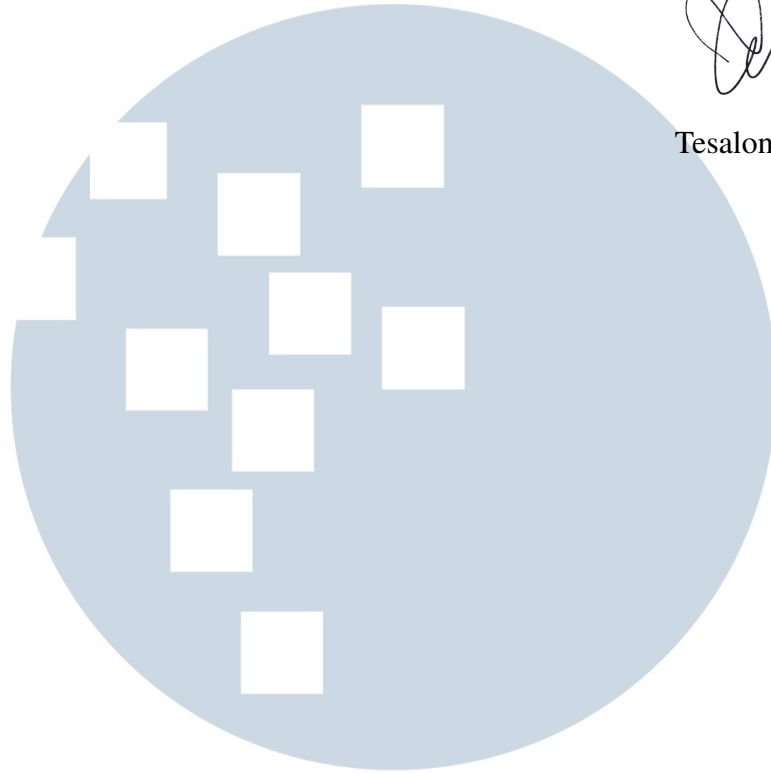
1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Yaman Khaeruzzaman, M.Sc., selaku PJS Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya skripsi ini.
5. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan moral, sehingga penulis dapat menyelesaikan skripsi ini.
6. dr. Jesslyn Valentina, M.M. yang telah bersedia menjadi ahli dan/atau narasumber dan meluangkan waktunya untuk melakukan *interview* untuk memenuhi kebutuhan dan kelengkapan dalam pengerjaan skripsi ini.
7. Teman-teman sekolah, kuliah, dan kantor yang telah memberikan bantuan dukungan moral selama pengerjaan skripsi, sehingga penulis dapat menyelesaikan skripsi ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 07 Mei 2024



Tesalonika Abigail



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

RANCANG BANGUN SISTEM KLASIFIKASI TINGKAT OBESITAS DENGAN ALGORITMA *RANDOM FOREST CLASSIFIER*

Tesalonika Abigail

ABSTRAK

Obesitas merupakan salah satu masalah kesehatan dunia serius yang disebabkan oleh berlebihnya timbunan lemak pada tubuh. Menurut *World Health Organization* (WHO), prevalensi obesitas telah meningkat 3 kali lipat dalam 40 tahun terakhir, dengan 650 juta dari 1.9 miliar orang dewasa yang kelebihan berat badan mengalami obesitas. Obesitas merupakan penyakit tidak menular yang berisiko menyebabkan penyakit lain yang lebih berbahaya seperti penyakit jantung dan kanker. Oleh karena itu, deteksi dini tingkat obesitas sangat penting. Saat ini, telah terdapat indikator pengukuran *Body Mass Index* (BMI), namun diketahui telah melebih-lebihkan obesitas bagi yang memiliki massa otot yang tinggi dan sebaliknya. Oleh karena itu, BMI dianggap kurang efektif karena hanya bergantung pada tinggi dan berat badan, tanpa mempertimbangkan komposisi tubuh dan keseharian seseorang. Untuk mengatasi permasalahan ini, telah dibangun model *Random Forest Classifier* terbaik yang dipilih berdasarkan hasil *model selection* setelah melakukan perbandingan dengan menerapkan *feature selection* dan *hyperparameter tuning*. Model yang terpilih berhasil meningkatkan akurasi model sebesar 1.4% yang kemudian diimplementasikan ke dalam sebuah sistem untuk mengklasifikasikan tingkat obesitas. Evaluasi model menghasilkan nilai *Precision*, *Recall*, *F1-Score*, dan akurasi sebesar 97%, 97%, 97%, dan 96.8% secara berurutan. Dari hasil evaluasi tersebut, dapat disimpulkan bahwa sistem ini mampu mengklasifikasikan tingkat obesitas dengan sangat baik.

Kata kunci: *Feature Selection*, *Hyperparameter Tuning*, Klasifikasi Tingkat Obesitas, *Model Selection*, *Random Forest Classifier*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Obesity Level Classification System with Random Forest Classifier Algorithm

Tesalonika Abigail

ABSTRACT (English)

Obesity is one of the serious global health problems caused by excessive accumulation of body fat. According to the World Health Organization (WHO), the prevalence of obesity has tripled in the last 40 years, with 650 million out of 1.9 billion overweight adults suffering from obesity. Obesity is a non-communicable disease that increases the risks of more dangerous diseases, such as heart disease and cancer. Therefore, early detection of obesity level is crucial. Currently, Body Mass Index (BMI) serves as a measurement indicator, but it tends to overestimate obesity for those with high muscle mass and vice versa, making it ineffective as it only relies on height and weight, without considering body composition and daily activities. To solve this, the best Random Forest Classifier model has been developed, selected based on the results of model selection after comparisons using feature selection and hyperparameter tuning. The selected model successfully improved accuracy by 1.4%, which then implemented into a system to classify obesity levels. Evaluation of the model resulted in Precision, Recall, F1-Score, and Accuracy of 97%, 97%, 97%, and 96.8% respectively. Based on these evaluation results, it can be concluded that this system is highly effective in classifying obesity levels.

Keywords: *Feature Selection, Hyperparameter Tuning, Model Selection, Obesity Level Classification, Random Forest Classifier*

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/ MOTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Obesitas	5
2.2 Pembelajaran Mesin	5
2.2.1 <i>Supervised Learning</i>	6
2.2.2 <i>Unsupervised Learning</i>	6
2.2.3 <i>Reinforcement Learning</i>	7
2.3 <i>Decision Tree</i>	7
2.4 <i>Random Forest</i>	8
2.5 <i>Randomized Search</i>	10
2.6 <i>Confusion Matrix</i>	11
BAB 3 METODOLOGI PENELITIAN	14
3.1 Pengumpulan Data	15
3.2 <i>Pre-processing</i>	17
3.3 Pembangunan Model	18
3.4 <i>Feature Selection</i> dan Pembangunan Model	18
3.5 <i>Hyperparameter Tuning</i> dan Pembangunan Model	19
3.6 Evaluasi Kinerja Model	20
3.7 Perancangan Sistem	21
3.7.1 <i>Flowchart</i>	22
3.7.2 <i>Prototype</i>	23
3.8 Implementasi Tampilan Antarmuka Pengguna	33
3.9 Implementasi Kode untuk Sistem	34
BAB 4 HASIL DAN DISKUSI	35
4.1 Spesifikasi Sistem	35
4.2 Implementasi dan Evaluasi Model	36
4.2.1 Pengumpulan Data	36
4.2.2 <i>Pre-processing</i>	36
4.2.3 Pembangunan Model Pertama	37
4.2.4 <i>Feature Selection</i> dan Pembangunan Model Kedua	38

4.2.5	<i>Hyperparameter Tuning</i> dan Pembangun Model Ketiga . . .	41
4.2.6	Evaluasi Kinerja Model	44
4.3	Implementasi Sistem	49
4.3.1	Tampilan Antarmuka Pengguna	49
4.3.2	Penerapan Kode untuk Sistem	67
BAB 5	SIMPULAN DAN SARAN	74
5.1	Simpulan	74
5.2	Saran	74
DAFTAR PUSTAKA	76



DAFTAR GAMBAR

Gambar 2.1	Struktur <i>Decision Tree</i> [1]	8
Gambar 2.2	Cara Kerja Algoritma <i>Random Forest</i> [2]	9
Gambar 3.1	<i>Framework</i> Penelitian	14
Gambar 3.2	<i>Detail</i> Tahapan untuk Proses Pengumpulan Data	15
Gambar 3.3	<i>Detail</i> Tahapan untuk Proses <i>Pre-processing</i>	17
Gambar 3.4	<i>Detail</i> Tahapan untuk Proses Pembangunan Model	18
Gambar 3.5	<i>Detail</i> Tahapan untuk Proses <i>Feature Selection</i> dan Pembangunan Model	19
Gambar 3.6	<i>Detail</i> Tahapan untuk Proses <i>Hyperparameter Tuning</i> dan Pembangunan Model	19
Gambar 3.7	<i>Detail</i> Tahapan untuk Proses Evaluasi Kinerja Model	21
Gambar 3.8	<i>Detail</i> Tahapan untuk Proses Perancangan Sistem	21
Gambar 3.9	Alur Sistem	22
Gambar 3.10	Tampilan untuk Halaman Data Diri	24
Gambar 3.11	Opsi untuk Pertanyaan 1 pada Halaman 1	24
Gambar 3.12	Tampilan untuk Validasi Input Halaman 1	25
Gambar 3.13	Tampilan untuk Halaman Kesehatan Keluarga dan Pola Makan	26
Gambar 3.14	Opsi untuk Pertanyaan 1 dan 2 pada Halaman 2	26
Gambar 3.15	Opsi untuk Pertanyaan 3 pada Halaman 2	26
Gambar 3.16	Opsi untuk Pertanyaan 4 pada Halaman 2	27
Gambar 3.17	Tampilan untuk Validasi Input Halaman 2	27
Gambar 3.18	Tampilan untuk Halaman Rutinitas	28
Gambar 3.19	Opsi untuk Pertanyaan 1 dan 7 pada Halaman 3	29
Gambar 3.20	Opsi untuk Pertanyaan 2 dan 4 pada Halaman 3	29
Gambar 3.21	Opsi untuk Pertanyaan 3 pada Halaman 3	29
Gambar 3.22	Opsi untuk Pertanyaan 5 pada Halaman 3	30
Gambar 3.23	Opsi untuk Pertanyaan 6 pada Halaman 3	30
Gambar 3.24	Opsi untuk Pertanyaan 8 pada Halaman 3	30
Gambar 3.25	Tampilan untuk Validasi Input Halaman 3	31
Gambar 3.26	Tampilan untuk Modalitas <i>Error</i>	32
Gambar 3.27	Tampilan untuk Modalitas Sukses	33
Gambar 3.28	<i>Detail</i> Tahapan untuk Proses Implementasi Tampilan Antarmuka Pengguna	33
Gambar 3.29	<i>Detail</i> Tahapan untuk Proses Implementasi Kode untuk Sistem	34
Gambar 4.1	Penggunaan <i>Dataset</i> dari Kaggle.com	36
Gambar 4.2	<i>Preprocessing Null & Duplicated Values</i>	37
Gambar 4.3	<i>Preprocessing Label Encoding</i>	37
Gambar 4.4	<i>Pembangunan Model Pertama dan Pengecekan Akurasi</i>	38
Gambar 4.5	Kodingan untuk <i>Feature Selection</i> dan Pembangunan Model Kedua	39
Gambar 4.6	Kodingan untuk Ringkasan Nilai Akurasi pada setiap Iterasi	40
Gambar 4.7	Fitur-Fitur Terpilih dan Pembangunan Model Kedua serta Pengecekan Akurasi dari Hasil <i>Feature Selection</i>	41

Gambar 4.8	<i>Hyperparameter Tuning with Randomized Search</i>	42
Gambar 4.9	Pembuatan Model dan Pengecekan Akurasi dari Hasil <i>Hyperparameter Tuning</i>	43
Gambar 4.10	Hasil Evaluasi <i>Confusion Matrix</i>	44
Gambar 4.11	Ringkasan Hasil TP, FN, FP, dan TN untuk setiap Kelas . .	45
Gambar 4.12	Proses Mendapatkan <i>Instance</i> untuk Kelas-Kelas "Normal_Weight" yang Salah Diklasifikasikan sebagai "Overweight_Level_I"	46
Gambar 4.13	Pengecekan Penyebab Tingginya Kesalahan Klasifikasi pada Kelas "Normal_Weight" yang Salah Diklasifikasikan sebagai "Overweight_Level_I"	47
Gambar 4.14	Hasil <i>Classification Report</i>	48
Gambar 4.15	Hasil Perhitungan <i>Micro Average</i>	48
Gambar 4.16	Hasil <i>Log Loss</i>	49
Gambar 4.17	Tampilan untuk Hasil Implementasi Halaman 1	50
Gambar 4.18	Opsi untuk Pertanyaan 1 pada Hasil Implementasi Halaman 1	51
Gambar 4.19	<i>Value</i> untuk Pertanyaan 1 pada Hasil Implementasi Halaman 1	51
Gambar 4.20	Tampilan untuk Validasi Input Hasil Implementasi Halaman 1	52
Gambar 4.21	Tampilan untuk Validasi Ketidaksesuaian <i>Format Value</i> untuk Pertanyaan 4 Hasil Implementasi Halaman 1	52
Gambar 4.22	Tampilan untuk Hasil Implementasi Halaman 2	53
Gambar 4.23	Opsi untuk Pertanyaan 1 dan 2 pada Hasil Implementasi Halaman 2	53
Gambar 4.24	<i>Value</i> untuk Pertanyaan 1 pada Hasil Implementasi Halaman 2	54
Gambar 4.25	<i>Value</i> untuk Pertanyaan 2 pada Hasil Implementasi Halaman 2	54
Gambar 4.26	Opsi untuk Pertanyaan 3 pada Hasil Implementasi Halaman 2	54
Gambar 4.27	<i>Value</i> untuk Pertanyaan 3 pada Hasil Implementasi Halaman 2	55
Gambar 4.28	Info Tambahan untuk Pertanyaan 3 pada Hasil Implementasi Halaman 2	56
Gambar 4.29	Opsi untuk Pertanyaan 4 pada Hasil Implementasi Halaman 2	56
Gambar 4.30	<i>Value</i> untuk Pertanyaan 4 pada Hasil Implementasi Halaman 2	57
Gambar 4.31	Tampilan untuk Hasil Implementasi Halaman 3	58
Gambar 4.32	Opsi untuk Pertanyaan 1 pada Hasil Implementasi Halaman 3	58
Gambar 4.33	<i>Value</i> untuk Pertanyaan 1 pada Hasil Implementasi Halaman 3	59
Gambar 4.34	Info Tambahan untuk Pertanyaan 1 pada Hasil Implementasi Halaman 3	59

Gambar 4.35	Opsi untuk Pertanyaan 2 pada Hasil Implementasi Halaman 3	60
Gambar 4.36	<i>Value</i> untuk Pertanyaan 2 pada Hasil Implementasi Halaman 3	60
Gambar 4.37	Opsi untuk Pertanyaan 3 pada Hasil Implementasi Halaman 3	60
Gambar 4.38	<i>Value</i> untuk Pertanyaan 3 pada Hasil Implementasi Halaman 3	61
Gambar 4.39	Info Tambahan untuk Pertanyaan 3 pada Hasil Implementasi Halaman 3	61
Gambar 4.40	Opsi untuk Pertanyaan 4 pada Hasil Implementasi Halaman 3	62
Gambar 4.41	<i>Value</i> untuk Pertanyaan 4 pada Hasil Implementasi Halaman 3	62
Gambar 4.42	Tampilan untuk Modalitas <i>Error Field</i> pada Hasil Implementasi	63
Gambar 4.43	Tampilan untuk Modalitas <i>Error</i> Ketidaksesuaian Format <i>Value</i> pada Hasil Implementasi	63
Gambar 4.44	Tampilan untuk Modalitas <i>Error Server</i> pada Hasil Implementasi	64
Gambar 4.45	Tampilan untuk Modalitas Sukses pada Hasil Implementasi	64
Gambar 4.46	Tampilan Klasifikasi untuk Keenam Kelas Lainnya	65
Gambar 4.47	Pengecekan <i>Required Fields</i>	68
Gambar 4.48	Pengecekan <i>Format Value</i> pada <i>Field</i> Berat Badan	69
Gambar 4.49	Penyimpanan Inputan dari Pengguna	69
Gambar 4.50	Pemanggilan API untuk Mendapatkan Hasil Klasifikasi	70
Gambar 4.51	Penyimpanan Model dengan Pickle	71
Gambar 4.52	Pesan Informasi Model Berhasil Tersimpan	71
Gambar 4.53	Pemanggilan API untuk Menyimpan Model	72
Gambar 4.54	Penerimaan dan Penyimpanan Data Inputan dari <i>Frontend</i>	72
Gambar 4.55	Proses Klasifikasi menggunakan Model yang sudah Disimpan	73

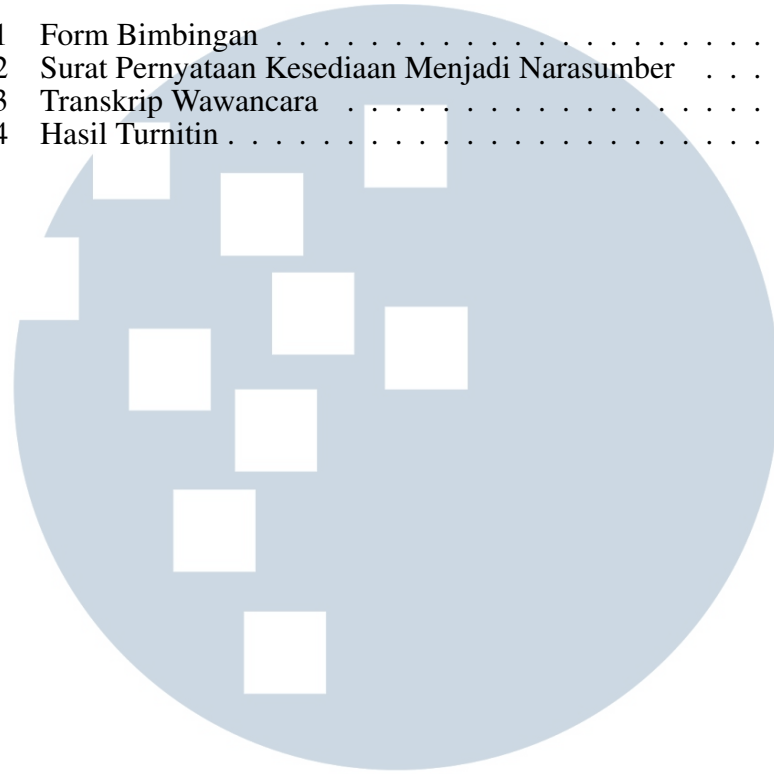
DAFTAR TABEL

Tabel 2.1	<i>Confusion Matrix</i> untuk <i>Binary Classification</i>	12
Tabel 2.2	<i>Confusion Matrix</i> untuk <i>Multi-class Classification</i>	13
Tabel 3.1	Pertanyaan dan Kemungkinan Jawaban dari Survei	16
Tabel 3.2	<i>Hyperparameter tuning</i> untuk pembuatan model <i>Random Forest</i>	20
Tabel 4.1	<i>Hyperparameter tuning</i> untuk pembangunan model ke-3	42
Tabel 4.2	Ringkasan Nilai Akurasi untuk Ketiga Model	43
Tabel 4.3	Nilai TP, TN, FP, dan FN untuk setiap Kelas	45
Tabel 4.4	Saran untuk Setiap Tingkatan Obesitas	66



DAFTAR LAMPIRAN

Lampiran 1	Form Bimbingan	79
Lampiran 2	Surat Pernyataan Kesediaan Menjadi Narasumber	81
Lampiran 3	Transkrip Wawancara	82
Lampiran 4	Hasil Turnitin	87



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA