

**IMPACT OF SMOTE AND ADASYN ON CLASS IMBALANCE
IN METABOLIC SYNDROME CLASSIFICATION USING
RANDOM FOREST ALGORITHM**

LAPORAN NON-REGULER SCIENTIST

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



Disusun oleh :

LUTFIANA DEKA NURHAYATI

22.11.4975

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

**IMPACT OF SMOTE AND ADASYN ON CLASS IMBALANCE IN
METABOLIC SYNDROME CLASSIFICATION USING RANDOM
FOREST ALGORITHM**

LAPORAN NON-REGULER SCIENTIST

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



Disusun oleh :

LUTFIANA DEKA NURHAYATI

22.11.4975

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

JALUR NON-REGULER SCIENTIST

**IMPACT OF SMOTE AND ADASYN ON CLASS IMBALANCE IN
METABOLIC SYNDROME CLASSIFICATION USING RANDOM FOREST
ALGORITHM**

yang disusun dan diajukan oleh

Lutfiana Deka Nurhayati

22.11.4975

telah disetujui oleh Dosen Pembimbing
pada tanggal 18 November 2025

Dosen Pembimbing,



Majid Rahardi, S.Kom, M.Eng
NIK. 190301393

HALAMAN PENGESAHAN

JALUR NON-REGULER SCIENTIST

**IMPACT OF SMOTE AND ADASYN ON CLASS IMBALANCE IN
METABOLIC SYNDROME CLASSIFICATION USING RANDOM FOREST
ALGORITHM**

yang disusun dan diajukan oleh

Lutfiana Deka Nurhayati
22.11.4975

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 November 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

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

Theophilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 November 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Lutfiana Deka Nurhayati

NIM : 22.11.4975

Menyatakan bahwa Laporan dengan judul berikut:

Impact of SMOTE and ADASYN on Class Imbalance in Metabolic Syndrome Classification Using Random Forest Algorithm

Dosen Pembimbing : Majid Rahardi, S.Kom, M.Eng

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, 18 November 2025

Yang Menyatakan,



Lutfiana Deka Nurhayati

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT atas segala rahmat dan petunjuk-Nya, karya ini penulis persembahkan kepada kedua orang tua dan adik tercinta yang selalu memberikan kasih sayang, doa, serta dukungan, kepada keluarga besar penulis yang senantiasa memberikan dukungan dan doa, serta kepada teman-teman yang penulis sayangi. Penulis berharap karya ini dapat memberikan manfaat bagi pembaca dan menjadi langkah awal yang kokoh serta bermanfaat bagi masa depan penulis.



KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa atas rahmat, karunia serta petunjuk-Nya sehingga penulis dapat menyelesaikan laporan yang berjudul "Impact of SMOTE and ADASYN on Class Imbalance in Metabolic Syndrome Classification Using Random Forest Algorithm". 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 Majid Rahardi, S.Kom, M.Eng, 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. Kusri, 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 Sipin dan Ibu Junariah, kedua orang tua penulis yang senantiasa memberikan doa, kasih sayang, semangat, dan dukungan baik moral maupun material.
7. Kepada Om Dhimas beserta istri, selaku keluarga penulis yang telah memberikan bantuan, dukungan serta kepercayaan yang sangat berarti dalam mendukung kelancaran studi penulis.
8. Kepada keluarga besar penulis yang telah memberikan dukungan dan semangat serta nasihat berharga.

9. Adik tercinta, Shinta Nur Ria yang menjadi sumber keceriaan dan kekuatan bagi penulis.
10. Teman-teman seperjuangan yang penulis sayangi Esti, Nisa, Vivi, serta Suryani 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.
11. Keluarga penjaga kos Putri 26, yang telah memberikan kenyamanan dan kebaikan selama masa studi penulis sehingga membuat penulis merasa seperti bersama keluarga.

Penulis menyadari bahwa laporan ini masih jauh dari kata sempurna karena keterbatasan penulis. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf atas segala kekurangan dan sangat mengharapkan kritik serta saran guna perbaikan dimasa mendatang.

Akhir kata, penulis berharap laporan ini dapat memberikan manfaat bagi pembaca dan bagi perkembangan ilmu pengetahuan, khususnya di bidang *Machine Learning*.

Yogyakarta, 19 November 2025

Penulis

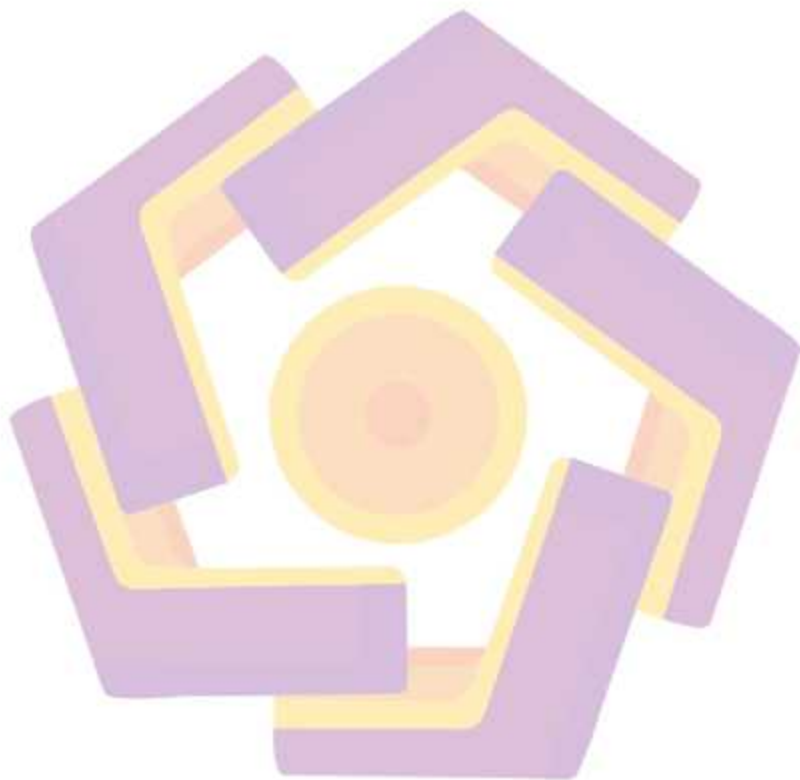
DAFTAR ISI

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	xvi
<i>Abstract</i>	xvii
Bab I Pendahuluan	1
1.1. Gambaran Umum.....	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	3
Bab II Tinjauan Pustaka	4
2.1. Studi Literatur	4
2.2. Landasan Teori.....	5
2.2.1 Sindrom Metabolik	5
2.2.2 Teknik Oversampling	5
2.2.3 Algoritma Random Forest.....	7
2.2.4 Confusion Matrix	8
2.2.5 Kurva ROC AUC.....	9
BAB III Metode Penelitian	10
3.1. Metode.....	10
3.1.1. Akuisisi Data.....	11

3.1.2.	Preprocessing	12
3.1.3.	Balancing Data	13
3.1.4.	Split Data.....	13
3.1.5.	Implementasi Model Random Forest.....	13
3.1.6.	Evaluasi Model.....	13
BAB IV Pembahasan		14
4.1.	Proses Balancing Data.....	14
4.2.	Uji Model dan Evaluasi.....	15
BAB V Kesimpulan		20
5.1.	Kesimpulan	20
5.2.	Saran.....	20
Referensi		22
Curriculum Vitae		25
Lampiran dan Bukti Pendukung.....		26
a.	Letter of Acceptance (LOA).....	26
b.	Lembar Review	27
c.	Bukti Terbit/Terindex	30
d.	Bukti pembayaran	30

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	8
Tabel 3. 1 Deskripsi Fitur Dataset	11
Tabel 4. 1 Distribusi Kelas Metabolic Syndrome	14
Tabel 4. 2 Evaluasi Performa	15



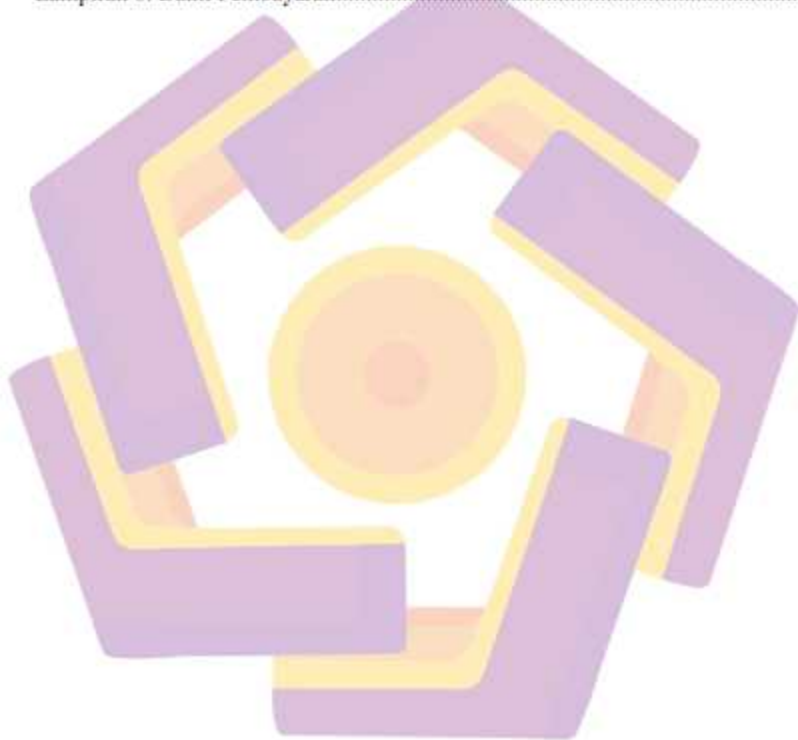
DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Algoritma Random Forest	8
Gambar 3. 1 Tahapan penelitian.....	10
Gambar 3. 2 Distribusi kelas target MetabolicSyndrome	12
Gambar 4. 1 Confusion Matrix Tanpa Penanganan Balancing	16
Gambar 4. 2 Confusion Matrix dengan Penanganan Balancing SMOTE.....	17
Gambar 4. 3 Confusion Matrix dengan Penanganan Balancing ADASYN.....	17
Gambar 4. 4 ROC-AUC.....	18



DAFTAR LAMPIRAN

Lampiran 1. Letter of Acceptance (LOA).....	26
Lampiran 2. Lembar Review.....	27
Lampiran 3. Lembar Review.....	28
Lampiran 4. Lembar Review.....	29
Lampiran 5. Terindex Sinta 3 Jurnal JAIC Vol. 9 No. 5 (2025); October 2025....	30
Lampiran 6. Bukti Pembayaran.....	30



DAFTAR LAMBANG DAN SINGKATAN

x_{syn}	Data baru hasil sintesis
x_i	Data yang akan direplikasi dari kelas minoritas yang dipilih acak
δ	Bilangan acak antara 0 dan 1
x_{knn}	Data terdekat dengan x_i dihitung dengan euclidean distance
r_i	Rasio ketidakseimbangan lokal sampel minoritas i
d_i	Jumlah tetangga mayoritas di sekitar sampel i
N_i	Jumlah total tetangga yang digunakan
g_i	Bobot dari data sintetis yang perlu dihasilkan sampel i
N_m	Jumlah total sampel minoritas
$\sum r_j$	Total rasio ketidakseimbangan dari semua sampel minoritas
G	Jumlah sampel sintetis yang akan dibuat
N_{maj}	Jumlah data kelas mayoritas
β	Parameter tingkat oversampling (antara 0 dan 1)
X_j	Tetangga dari sampel i
SMOTE	<i>Synthetic Minority Over-sampling</i>
ADASYN	<i>Adaptive Synthetic Sampling</i>
BMI	<i>Body Mass Index</i>
HDL	<i>High Density Lipoprotein</i>
TP	<i>True Positive</i>
TN	<i>True Negatif</i>
FP	<i>False Positive</i>
FN	<i>False Negatif</i>
ROC	<i>Receiver Operating Characteristic</i>
AUC	<i>Area Under the Curve</i>
TPR	<i>True Positif Rasio</i>
FPR	<i>False Positive Rasio</i>
ROS	<i>Random Over Sampling</i>
RUS	<i>Random Under Sampling</i>

DAFTAR ISTILAH



Metabolic Syndrome	kelainan medis yang meliputi kondisi obesitas sentral, resistensi insulin, hipertensi dan dislipidemia.
MetSyn	individu yang terdiagnosa Sindrom Metabolik.
NoMetSyn	individu yang tidak terdiagnosa Sindrom Metabolik.
Class imbalance	ketidakseimbangan kelas pada dataset.
Machine Learning	salah satu cabang dari kecerdasan buatan yang memungkinkan sistem untuk mempelajari pola berdasarkan data data untuk membuat prediksi atau keputusan secara otomatis.
Oversampling	teknik untuk meningkatkan jumlah sampel pada kelas minoritas sehingga jumlah antara kelas mayoritas dan kelas minoritas menjadi lebih seimbang.
SMOTE	metode oversampling yang bekerja dengan menciptakan data sintetis baru pada kelas minoritas
ADASYN	metode oversampling yang menambahkan sampel baru pada kelas minoritas yang sulit diklasifikasi.
Random Forest	algoritma klasifikasi yang bekerja dengan membangun beberapa pohon keputusan dan menggabungkan hasil prediksi melalui pemungutan suara.
Overfitting	kondisi ketika model terlalu sempurna pada data pelatihan, sehingga kehilangan kemampuan untuk memprediksi data baru secara akurat.
Label Encoding	teknik mengubah data kategorikal menjadi data numerik.
Feature Selection	proses seleksi fitur yang relevan.



Confusion Matrix	metode evaluasi yang digunakan dalam machine learning untuk mengukur kinerja dari sebuah model klasifikasi.
Akurasi	tingkat keakuratan sebuah model.
Presisi	tingkat ketepatan model dalam memprediksi kelas yang diminta.
Recall	tingkat sensitivitas model.
F1-Score	tingkat keseimbangan antara presisi dan recall.
Specificity	tingkat ketepatan model dalam memprediksi kelas negatif.
ROC	untuk mengevaluasi performa klasifikasi secara visual.
AUC	nilai dari area dibawah kurva.
True Positive	data positif diprediksi benar.
True Negatif	data negatif diprediksi benar.
False Positive	data negatif salah diprediksi sebagai positif.
False Negatif	data positif salah diprediksi sebagai negatif.

INTISARI

Sindrom Metabolik merupakan kumpulan kondisi medis yang dapat meningkatkan risiko terkena stroke, penyakit kardiovaskular, dan diabetes tipe 2. Deteksi dini terhadap kondisi ini memerlukan model *machine learning* yang mampu melakukan klasifikasi secara akurat guna mendukung penanganan tepat waktu. Namun, terkadang ketidakseimbangan kelas pada data seringkali mempengaruhi kinerja algoritma klasifikasi, terutama dalam mengenali kelas minoritas yaitu individu terdiagnosis Sindrom Metabolik. Penelitian ini bertujuan untuk menganalisis pengaruh teknik SMOTE dan ADASYN dalam menangani ketidakseimbangan data pada klasifikasi Sindrom Metabolik menggunakan algoritma Random Forest. Algoritma tersebut dipilih karena kemampuannya menghasilkan prediksi yang akurat, meskipun performanya dapat menurun ketika dihadapkan pada distribusi kelas yang tidak seimbang. Hasil penelitian menunjukkan bahwa model tanpa teknik penyeimbangan data mencapai akurasi sebesar 86% dengan recall pada kelas minoritas sebesar 75%. Penerapan SMOTE meningkatkan akurasi menjadi 91% dan recall 93%, sedangkan ADASYN memperoleh akurasi 92% dengan recall kelas minoritas sebesar 95%. Temuan ini mengindikasikan bahwa teknik ADASYN yang dikombinasikan dengan algoritma Random Forest memberikan peningkatan performa yang signifikan dalam klasifikasi Sindrom Metabolik pada data yang tidak seimbang.

Kata kunci: ADASYN, Sindrom Metabolik, SMOTE, Random Forest.

ABSTRACT

Metabolic Syndrome is a collection of medical conditions that can increase the risk of stroke, cardiovascular disease, and type 2 diabetes. Early detection of this condition requires a machine learning model capable of accurate classification to support timely treatment. However, class imbalance in data often hampers the performance of classification algorithms, particularly in recognizing minority classes, namely individuals diagnosed with Metabolic Syndrome. This study aims to analyze the effect of applying the SMOTE and ADASYN data balancing techniques in classifying Metabolic Syndrome using the Random Forest algorithm. These algorithms were chosen for their ability to produce accurate predictions, although their performance can decline when faced with imbalanced class distributions. The results showed that the model without data balancing techniques achieved 86% accuracy with a minority class recall of 75%. The application of SMOTE increased accuracy to 91% and recall to 93%, while ADASYN achieved 92% accuracy and a minority class recall of 95%. These findings indicate that the ADASYN technique combined with the Random Forest algorithm provides significant performance improvements in the classification of Metabolic Syndrome on imbalanced data.

Keyword: ADASYN, Metabolic Syndrome, SMOTE, Random Forest.