



KARYA TULIS ILMIAH

*Heart-Risk Classifier : Klasifikasi dan
Feature Importance Penyakit Jantung
Berbasis Machine Learning (ML)
Menggunakan Data Mikro
SKI 2023*

ZAHRA MULKI SYARI'ATI
NIM.P2.06.37.1.23.078

PROGRAM STUDI DIPLOMA TIGA RMIK CIREBON

JURUSAN REKAM MEDIS DAN INFORMASI KESEHATAN
POLITEKNIK KESEHATAN TASIKMALAYA
KEMENTERIAN KESEHATAN REPUBLIK INDONESIA
TAHUN 2026



KARYA TULIS ILMIAH

Diajukan Dalam Rangka Memenuhi Salah Satu Syarat Menyelesaikan Pendidikan D-III Program
Studi Rekam Medis dan Informasi Kesehatan

Heart-Risk Classifier : Klasifikasi dan Feature Importance Penyakit Jantung Berbasis Machine Learning (ML) Menggunakan Data Mikro SKI 2023

**ZAHRA MULKI SYARI'ATI
NIM.P2.06.37.1.23.078**

PROGRAM STUDI DIPLOMA TIGA RMIK CIREBON

**JURUSAN REKAM MEDIS DAN INFORMASI KESEHATAN
POLITEKNIK KESEHATAN TASIKMALAYA
KEMENTERIAN KESEHATAN REPUBLIK INDONESIA
TAHUN 2026**

LEMBAR PERNYATAAN ORISINALITAS

Karya Tulis Ilmiah (KTI) ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan sendiri dengan benar.

Nama : Zahra Mulki Syari'ati

NIM : P2.06.37.1.23.078

Tanggal : 8 Mei 2026

Tanda Tangan :



**LEMBAR PERNYATAAN
PERSETUJUAN PUBLIKASI**

Saya yang bertanda tangan dibawah ini

Nama : Zahra Mulki Syari'ati

NIM : P2.06.37.1.23.078

Program Studi : D-III Rekam Medis dan Informasi Kesehatan Cirebon

Judul Karya Tulis : *Heart-Risk Classifier: Klasifikasi dan Feature importance*
Penyakit Jantung Berbasis *Machine Learning* (ML)
Menggunakan Data Mikro SKI 2023

Menyatakan dengan sebenarnya bahwa saya mengizinkan/menyetujui Karya Tulis Ilmiah yang saya susun dipublikasikan untuk kepentingan akademis, baik secara keseluruhan maupun sebagian dari karya tulis ini.

Demikian pernyataan ini dibuat untuk dapat dipergunakan sebagaimana mestinya.

Cirebon, 8 Mei 2026

Yang Membuat Pernyataan



Zahra Mulki Syari'ati

NIM. P2.06.37.1.23.078

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN KARYA TULIS ILMIAH

HEART-RISK CLASSIFIER: KLASIFIKASI DAN FEATURE IMPORTANCE ***PENYAKIT JANTUNG BERBASIS MACHINE LEARNING (ML)*** ***MENGGUNAKAN DATA MIKRO SKI 2023***

Dipersiapkan dan disusun oleh:

ZAHRA MULKI SYARI'ATI
P2.06.37.1.23.078

Telah dipertahankan dihadapan Tim Penguji

pada Tanggal

8 Mei 2026

Ketua Penguji

Maula Ismail M. S.T. M.KM
NIP. 198212282006041007

Penguji I

Lina Khasanah, S.KM, M.KM
NIP. 198108092009122002

Penguji II

Bambang Karmanto, S.KM, M.Kes
NIP. 196304211986031003

Mengetahui,
Ketua Program Studi,



iv

UNGKAPAN TERIMA KASIH

Puji syukur kehadiran Allah SWT atas limpahan rahmat, taufik, dan hidayah-Nya, penulis dapat menyelesaikan Karya Tulis Ilmiah (KTI) yang berjudul “*Heart-Risk Classifier: Klasifikasi dan Feature importance Penyakit Jantung Berbasis Machine Learning (ML) Menggunakan Data Mikro SKI 2023*”, dengan tepat waktu dan sebaik-baiknya.

Penyusunan Karya Tulis Ilmiah (KTI) ini ditujukan sebagai salah satu syarat untuk menyelesaikan pendidikan Diploma III Rekam Medis dan Informasi Kesehatan di Politeknik Kesehatan Kemenkes Tasikmalaya. Penulis menyadari bahwa tanpa adanya bantuan, dorongan, dan bimbingan dari berbagai pihak, karya tulis ini tidak mungkin dapat terselesaikan dengan baik. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada:

1. Badan Kebijakan Pembangunan Kesehatan (BKPK) Kementerian Kesehatan Republik Indonesia, selaku penyedia data Survei Kesehatan Indonesia (SKI) 2023 yang menjadi dasar penelitian ini;
2. Ibu Dr. Dini Mariani, S.Kep., Ners, M.Kep., selaku Direktur Politeknik Kesehatan Kemenkes Tasikmalaya;
3. Bapak Andi Suhenda, S.KM, MPH., selaku Ketua Jurusan Rekam Medis dan Informasi Kesehatan Politeknik Kesehatan Kemenkes Tasikmalaya;
4. Ibu Elfi, S.ST., MPH., selaku Ketua Program Studi Rekam Medis dan Informasi Kesehatan Cirebon Politeknik Kesehatan Kemenkes Tasikmalaya;
5. Bapak Maula Ismail Muhammad, S.T., M.KM., selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan arahan selama proses penyusunan karya tulis ini;
6. Ibu Lina Khasanah, S.KM, M.KM dan Bapak Bambang Karmanto, S.KM, M.Kes, selaku dosen penguji yang telah memberikan masukan berharga dalam penyempurnaan karya tulis ini;
7. Civitas akademika Prodi D-III Rekam Medis dan Informasi Kesehatan Cirebon yang telah memberikan ilmu dan pengalaman berharga kepada penulis;

8. Kedua orang tua tercinta, yang doanya menjadi nafas setiap langkah, yang kasih sayangnya menjadi cahaya dalam kegelapan, yang kesabarannya menjadi alasan penulis terus berjuang. Tanpa doa, pelukan, dan pengorbanan yang tak terlihat, perjalanan ini tidak akan pernah sampai pada titik ini;
9. Teman-teman seperjuangan penulis yang selalu hadir sebagai teman perjalanan dalam belajar dan berkembang. Terima kasih atas kebersamaan dan dukungan yang membuat proses ini terasa lebih ringan dijalani;
10. Semua pihak yang tidak dapat disebutkan satu per satu, namun telah memberikan kontribusi dalam penyelesaian karya tulis ini;

Akhir kata, penulis berharap semoga karya tulis ilmiah ini dapat memberikan manfaat dan menjadi sumbangan kecil dalam pengembangan ilmu pengetahuan di bidang rekam medis dan informasi kesehatan.

Cirebon, Mei 2026

Penulis

**Kementerian Kesehatan Republik Indonesia
Politeknik Kesehatan Tasikmalaya
Jurusan Rekam Medis dan Informasi Kesehatan
Program Studi Diploma III RMIK Cirebon
2026**

ZAHRA MULKI SYARI'ATI

***HEART-RISK CLASSIFIER: KLASIFIKASI DAN FEATURE IMPORTANCE
PENYAKIT JANTUNG BERBASIS MACHINE LEARNING (ML)
MENGUNAKAN DATA MIKRO SKI 2023***
68 Halaman, 5 Bab, 2 Tabel, 6 Gambar, 13 Lampiran

ABSTRAK

Latar Belakang: Penyakit jantung merupakan penyebab kematian tertinggi di Indonesia dengan peningkatan angka kematian lebih dari 25% menurut laporan *Global Burden of Disease 2020*. Data *Institute for Health Metrics and Evaluation (IHME)* mencatat penyakit jantung menyumbang 53,2% kematian nasional. Dari sisi pembiayaan, klaim Jaminan Kesehatan Nasional (JKN) untuk penyakit jantung mencapai Rp17,92 triliun pada 2022, menjadikannya penyakit katastrofik dengan beban tertinggi. Survei Kesehatan Indonesia (SKI) 2023 yang melibatkan lebih dari 586.000 rumah tangga menyediakan data mikro komprehensif. Tujuan penelitian ini adalah membangun model klasifikasi risiko penyakit jantung berbasis data SKI 2023 serta mengidentifikasi faktor risiko utama dengan *Machine Learning (ML)*.

Metodologi Penelitian: Penelitian kuantitatif dengan desain *case-control* menggunakan data mikro SKI 2023. Populasi 602.982 responden dewasa, dengan sampel dipilih sesuai kriteria inklusi dan eksklusi. Variabel dependen adalah status penyakit jantung, sedangkan variabel independen meliputi usia, aktivitas fisik, konsumsi buah dan sayur, Indeks Massa Tubuh (IMT), riwayat hipertensi, dan merokok. Analisis dilakukan dengan kerangka *Knowledge Discovery in Database (KDD)* menggunakan algoritma *Random Forest* dan evaluasi *feature importance*.

Hasil: Model *Random Forest* memiliki performa klasifikasi yang baik berdasarkan indikator akurasi, *sensitivity*, dan *specificity*. Analisis *feature importance* menunjukkan bahwa variabel usia, riwayat hipertensi, dan Indeks Massa Tubuh (IMT) merupakan faktor risiko dominan dalam menentukan kejadian penyakit jantung. Sementara itu, variabel perilaku seperti merokok, aktivitas fisik, serta pola konsumsi berperan sebagai faktor pendukung yang memodulasi tingkat risiko.

Kesimpulan: Pendekatan *Machine Learning* menggunakan algoritma *Random Forest* efektif dalam mengklasifikasikan risiko penyakit jantung serta mampu mengidentifikasi faktor risiko utama secara komprehensif.

Kata kunci: Faktor Risiko, *Feature importance*, *Machine Learning (ML)*, Penyakit Jantung, SKI 2023

Daftar Pustaka: 83 (2021-2026)

ZAHRA MULKI SYARI'ATI

***HEART-RISK CLASSIFIER: CLASSIFICATION AND FEATURE
IMPORTANCE HEART DISEASE BASED ON MACHINE LEARNING (ML)
USING MIKRO SKI 2023 DATA
68 Pages, 5 Chapter, 2 Tables, 6 Image, 13 Attachments***

ABSTRACT

Background: Heart disease is the leading cause of death in Indonesia, with a mortality rate increase of more than 25% according to the Global Burden of Disease 2020 report. Data from the Institute for Health Metrics and Evaluation (IHME) shows that heart disease accounts for 53.2% of national deaths. In terms of financing, National Health Insurance (JKN) claims for heart disease reached IDR 17.92 trillion in 2022, making it the most catastrophic disease with the highest burden. The 2023 Indonesian Health Survey (SKI), which involved more than 586,000 households, provided comprehensive microdata. The purpose of this study was to build a data-based risk classification model for heart disease using the 2023 SKI data and to identify the main risk factors using Machine Learning (ML).

Research Method: A quantitative study with a case-control design using SKI 2023 microdata. The population consisted of 602,982 adult respondents, with samples selected according to inclusion and exclusion criteria. The dependent variable was heart disease status, while the independent variables included age, physical activity, fruit and vegetable consumption, Body Mass Index (BMI), history of hypertension, and smoking. The analysis was performed using the Knowledge Discovery in Database (KDD) framework with the Random Forest algorithm and feature importance evaluation.

Results: The Random Forest model demonstrated good classification performance based on accuracy, sensitivity, and specificity metrics. Feature importance analysis revealed that age, history of hypertension, and Body Mass Index (BMI) were the dominant risk factors in determining the occurrence of heart disease. Meanwhile, behavioral variables such as smoking, physical activity, and dietary patterns acted as supporting factors that modulate the level of risk.

Conclusion: The Machine Learning approach using the Random Forest algorithm is effective in classifying heart disease risk and is capable of comprehensively identifying the main risk factors.

Keywords: Risk Factors, Feature importance, Machine Learning (ML), Heart Disease, SKI 2023

Bibliography: 83 (2021-2026)

DAFTAR ISI

LEMBAR PERNYATAAN ORISINALITAS	Error! Bookmark not defined.
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	i
LEMBAR PERSETUJUAN KARYA TULIS ILMIAH.....	Error! Bookmark not defined.
LEMBAR PENGESAHAN KARYA TULIS ILMIAH.....	Error! Bookmark not defined.
UNGKAPAN TERIMA KASIH.....	iii
ABSTRAK	v
DAFTAR ISI.....	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN	xi
DAFTAR SINGKATAN.....	xii
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Rumusan Masalah	7
C. Tujuan Penelitian.....	7
D. Manfaat Penelitian	7
E. Keaslian Penelitian.....	8
BAB II TINJAUAN PUSTAKA.....	10
A. Tinjauan Pustaka	10
1. Penyakit Jantung	10
2. Faktor Risiko Penyakit Jantung	11
3. Survei Kesehatan Indonesia (SKI) 2023	14
4. <i>Machine Learning</i> (ML)	15
5. <i>Knowledge Discovery in Database</i> (KDD).....	20
B. Kerangka Teori	21
C. Kerangka Konseptual Penelitian	22
BAB III METODOLOGI PENELITIAN	24
A. Jenis dan Desain Penelitian	24
B. Tempat dan Waktu Penelitian.....	24
1. Tempat Penelitian.....	24

2. Waktu Penelitian	25
C. Populasi dan Sampel Penelitian	25
1. Populasi.....	25
2. Sampel.....	25
D. <i>Novelty</i> /Kebaruan Penelitian	26
E. Variabel Penelitian	26
1. Variabel Terikat (<i>Dependen</i>)	27
2. Variabel Bebas (<i>Independen</i>)	27
F. Definisi Operasional.....	28
G. Instrumen Penelitian dan Cara Pengumpulan Data.....	30
H. Teknik Analisis Data	31
I. Etika Penelitian	34
J. Keterbatasan Penelitian	35
K. Jalannya Penelitian.....	35
L. Jadwal Penelitian.....	36
BAB IV HASIL DAN PEMBAHASAN.....	38
A. Hasil	38
1. Gambaran Umum Data Mikro SKI 2023	38
2. Analisis Faktor Risiko Dominan Model	40
3. Evaluasi Kinerja Model <i>Machine Learning</i> (ML)	45
B. Pembahasan.....	47
1. Gambaran Umum Data Mikro SKI 2023	47
2. Analisis Faktor Risiko Dominan Model.....	50
3. Evaluasi Kinerja Model <i>Machine Learning</i>	54
BAB V KESIMPULAN DAN SARAN	57
A. Kesimpulan	57
B. Saran.....	58
DAFTAR PUSTAKA	59
LAMPIRAN.....	68

DAFTAR TABEL

Tabel 1.1 Keaslian Penelitian.....	8
Tabel 3.1 Definisi Operasional (Do) Variabel Penelitian.....	29
Tabel 3.2 Jadwal Penelitian.....	29

DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori Faktor Risiko Penyakit Jantung	21
Gambar 2.2 Kerangka Teori <i>Framework</i> KDD Dalam <i>Machine Learning</i> (ML).	22
Gambar 2.3 Kerangka Konseptual Penelitian	23
Gambar 3.1 Mekanisme <i>Framework</i> KDD.....	31
Gambar 4.1. Hasil Analisis <i>Feature importance</i>	40
Gambar 4.2 <i>Confusion Matrix Model Random Forest</i>	44

DAFTAR LAMPIRAN

Lampiran 1. Lembar <i>Checklist</i> Ketersediaan Data Penelitian	69
Lampiran 2. Pernyataan Penggunaan <i>Artificial Intelligence</i> (AI).....	70
Lampiran 3. Lembar Turnitin.....	71
Lampiran 4. Lembar Konsultasi Proposal Karya Tulis Ilmiah	72
Lampiran 5. Lembar Konsultasi Karya Tulis Ilmiah.....	74
Lampiran 6. Surat Permohonan Data Mikro SKI 2023.....	76
Lampiran 7. Pengolahan Data Mikro SKI 2023 dengan SPSS.....	77
Lampiran 8. <i>Variable View</i> Data Mikro SKI.....	77
Lampiran 9. <i>Worksheet Orange Data Mining</i>	78
Lampiran 10. <i>Data Table Orange Data Mining</i> dalam Variabel Penelitian.....	78
Lampiran 11. Edit Domain Variabel.....	79
Lampiran 12. Hasil <i>Feature importance</i>	79
Lampiran 13. Hasil Evaluasi Model <i>Random Forest</i>	80

DAFTAR SINGKATAN

ML	: <i>Machine Learning</i>
KDD	: <i>Knowledge Discovery in Database</i>
RME	: Rekam Medis Elektronik
SMOTE	: <i>Synthetic Minority Oversampling Technique</i>
PTM	: Penyakit Tidak Menular
IG	: <i>Information Gain</i>
GR	: <i>Gain Ratio</i>
SIK	: Sistem Informasi Kesehatan
SKI	: Survei Kesehatan Indonesia
SDGs	: <i>Sustainable Development Goals</i>