

SKRIPSI



**Implementasi Algoritma *Random Forest* untuk Klasifikasi
Risiko Diabetes Melitus Tipe 2 pada Data Pasien Puskesmas Palmerah**

RIVALDO ELSHADDAI KALONA SIANTURI

2110512063

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
JAKARTA
2026**

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Rivaldo Elshaddai Kalona Sianturi
NIM : 2110512063
Program Studi : S1 Sistem Informasi

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 22 Januari 2026

Yang Menyatakan,



Rivaldo Elshaddai Kalona Sianturi

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Rivaldo Elshaddai Kalona Sianturi
NIM : 2110512063
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

Demi Pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

Implementasi Algoritma *Random Forest* untuk Klasifikasi Risiko Diabetes Melitus Tipe 2 pada Data Pasien Puskesmas Palmerah

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Jakarta, 22 Januari 2026

Yang Menyatakan,



Rivaldo Elshaddai Kalona Sianturi

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Implementasi Algoritma Random Forest untuk Klasifikasi Risiko Diabetes
Melitus Tipe 2 pada Data Pasien Puskesmas Palmerah
Nama : Rivaldo Elshaddai Kalona Sianturi
NIM : 2110512063
Program Studi : SI Sistem Informasi

Disetujui oleh :

Penguji 1:

Iin Ernawati, S.Kom., M.Si.

Penguji 2:

Bambang Triwahyono, S.Kom., M.Si.

Pembimbing 1:

Ruth Mariana Bunga Wadu, S.Kom., MMSI.

Pembimbing 2:

Sarika, S.Kom., M.Kom

Diketahui oleh:

Koordinator Program Studi:

Anita Muliawati, S.Kom., MTI.

NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer:

Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM

NIP. 197605082003121002

Tanggal Ujian Tugas Akhir : 24 November 2025



ABSTRAK

Penelitian ini bertujuan untuk mengklasifikasikan tingkat risiko Diabetes Melitus Tipe 2 dengan memanfaatkan algoritma *Random Forest* menggunakan data pasien Puskesmas Palmerah tahun 2024. Data yang digunakan merupakan data rekam medis pasien yang telah melalui tahapan prapemrosesan, meliputi pembersihan data, transformasi variabel, dan pembagian data latih serta data uji. Model *Random Forest* dibangun dengan mengombinasikan sejumlah pohon keputusan untuk menghasilkan prediksi yang lebih stabil dan akurat. Evaluasi kinerja model dilakukan menggunakan *confusion matrix*, nilai akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mampu memberikan performa klasifikasi yang cukup baik dalam mengidentifikasi risiko Diabetes Melitus Tipe 2, sehingga diharapkan dapat membantu pihak puskesmas dalam mendukung pengambilan keputusan serta upaya pencegahan penyakit diabetes secara lebih dini.

Kata kunci: Diabetes Melitus Tipe 2, *Random Forest*, Klasifikasi.

ABSTRACT

This study aims to classify the risk level of Type 2 Diabetes Mellitus by applying the Random Forest algorithm using patient data from Palmerah Community Health Center in 2024. The dataset consists of patients' medical records that have undergone data preprocessing stages, including data cleaning, variable transformation, and the division of training and testing datasets. The Random Forest model was developed by combining multiple decision trees to produce more stable and accurate predictions. Model performance was evaluated using a confusion matrix, accuracy, precision, recall, and f1-score metrics. The results indicate that the Random Forest algorithm demonstrates good classification performance in identifying the risk of Type 2 Diabetes Mellitus. Therefore, this model is expected to support healthcare providers at the community health center in decision-making processes and in implementing early prevention strategies for diabetes.

Keywords: *Type 2 Diabetes Mellitus, Random Forest, Classification.*

KATA PENGANTAR

Puji dan Syukur saya panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunianya, sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan judul “Implementasi Algoritma *Random Forest* untuk Klasifikasi Risiko Diabetes Melitus Tipe 2 pada Data Pasien Puskesmas Palmerah”. Penulis menyadari dalam penyelesaian Tugas Akhir ini banyak pihak yang telah membimbing dan memberi dukungan. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih yang sebesar besarnya kepada yang terhormat :

1. Tuhan Yang Maha Esa atas rahmat dan karunia-Nya yang sungguh baik.
2. Alm. Ayah saya Rahmat Sianturi dan Ibu saya Katrin Ervina yang selalu memberi doa, dukungan, semangat dan motivasi tanpa henti dalam setiap langkah perjalanan penulis.
3. Pak Prof.. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku Dekan Fakultas Ilmu Komputer UPNVJ
4. Ibu Anita Muliawati, S.Kom. M.TI, S.Pd, M.Kom, selaku Koordinasi Program Studi S1 Sistem Informasi.
5. Ibu Ruth Mariana Bunga Wadu S.Kom., MMSI. selaku Dosen Pembimbing 1 dan Ibu Sarika, S.Kom., M.Kom. selaku Dosen Pembimbing 2, yang telah membimbing dan bersedia meluangkan waktu serta mengarahkan penulis dalam proses pengerjaan skripsi ini.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan di masa yang akan datang. Penulis berharap semoga skripsi ini dapat memberikan manfaat bagi semua pihak, serta dapat menjadi referensi untuk pengembangan ilmu pengetahuan di bidang Sistem Informasi, khususnya dalam analisis data Diabetes Melitus Tipe 2 menggunakan metode *Random Forest*.

DAFTAR ISI

PERNYATAAN ORISINALITAS.....	i
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	ii
LEMBAR PENGESAHAN.....	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL	x
DAFTAR SIMBOL	xi
BAB I.....	1
PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Luaran Penelitian.....	4
1.7. Sistematika Penulisan.....	4
BAB II.....	6
TINJAUAN PUSTAKA.....	6
2.1. Diabetes Melitus Tipe 2.....	6
2.2. Puskesmas.....	6
2.3. <i>Data Mining</i>	7
2.4. <i>Random Forest</i>	8
2.5. <i>Confusion Matrix</i>	11
2.6. <i>Streamlit</i>	12
2.7. Penelitian Terdahulu.....	13
BAB III.....	15
METODE PENELITIAN.....	15

3.1. Tahapan Penelitian.....	15
3.2. Studi Pustaka.....	16
3.3. Identifikasi Masalah.....	16
3.4. Pengumpulan Data.....	16
3.5. <i>Preprocessing Data</i>	16
3.6. <i>Exploratory Data Analysis (EDA)</i>	17
3.7. Pembagian Data Latih & Data Uji.....	17
3.8. Pembentukan Model.....	17
3.9. Evaluasi Model.....	17
3.10. Implementasi Sistem.....	18
3.11. Pengujian Sistem.....	18
3.12. Alat Bantu Penelitian.....	18
3.13. Jadwal Rencana Penelitian.....	19
BAB IV	20
HASIL & PEMBAHASAN	20
4.1. Persiapan Data.....	20
4.2. <i>Preprocessing Data</i>	23
4.3. <i>Exploration Data Analysis (EDA)</i>	24
4.4. Pembentukan Label Risiko Diabetes.....	26
4.5. Pembagian Data Latih & Data Uji.....	26
4.6. Pembentukan Model.....	27
4.7. Evaluasi Model.....	29
4.8. Implementasi dan Pengujian Sistem.....	31
BAB V.....	36
PENUTUP	36
5.1. Kesimpulan.....	36
5.2. Saran.....	37
DAFTAR PUSTAKA	38
DAFTAR RIWAYAT HIDUP.....	40
LAMPIRAN	41

DAFTAR GAMBAR





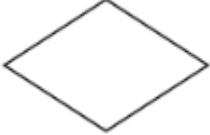
Gambar 2.1. Tahapan <i>Knowledge Discovery in Databases</i> (Course-Net, 2020).....	7
Gambar 3.1. Diagram Tahapan Penelitian	15
Gambar 4.1. Persiapan Dataset.....	23
Gambar 4.2. Dataset Setelah Menghapus <i>Missing Value</i>	23
Gambar 4.3. Dataset Setelah Penanganan Nilai <i>Outlier</i>	24
Gambar 4.4. Hasil Analisis Statistik Deskriptif.....	25
Gambar 4.5. Hasil Korelasi Antar Atribut.....	26
Gambar 4.6. Pembagian Data Latih & Data Uji.....	27
Gambar 4.7. Pelatihan Model <i>Random Forest</i>	28
Gambar 4.8. Proses Prediksi Status Diabetes Menggunakan Model <i>Random Forest</i>	28
Gambar 4.9. Evaluasi Model Random Forest.....	29
Gambar 4.10. <i>Feature Importance</i> Diabetes Model.....	30
Gambar 4.11. Halaman Beranda Sistem.....	31
Gambar 4.12. Halaman Beranda Sistem (Ringkasan Risiko Pasien).....	32
Gambar 4.13. Halaman Prediksi Pasien	33
Gambar 4.14. Tabel Hasil Prediksi Pasien.....	33
Gambar 4.15. Halaman Edukasi & FAQ Diabetes.....	34
Gambar 4.16. Halaman Tentang Sistem.....	35

DAFTAR TABEL

Tabel 2.1. Confusion Matrix.....	11
Tabel 2.2. Tabel Penelitian Terdahulu.....	13
Tabel 3.1 Jadwal Rencana Penelitian.....	19
Tabel 4.1. Daftar Atribut pada Dataset.....	20

DAFTAR SIMBOL

Simbol *Flowchart*

No	Simbol	Nama	Arti Simbol
1		<i>Terminator</i>	Menyatakan awal dan akhir dari sebuah program
2		<i>Process</i>	Menyatakan sebuah proses yang dilakukan oleh komputer
3		<i>Input/Output</i>	Simbol yang menyatakan proses input atau output tanpa tergantung dengan jenis peralatannya
4		<i>Flow</i>	Menggabungkan antara simbol satu dengan simbol lainnya
5		<i>Decision</i>	Menunjukkan kondisi tertentu yang menghasilkan dua kemungkinan jawaban, yaitu ya dan tidak