



**IMPLEMENTASI *PARTICLE SWARM OPTIMIZATION* PADA
ALGORITMA *NAIVE BAYES*, *DECISION TREE C4.5*, DAN *RANDOM
FOREST* UNTUK KLASIFIKASI DIABETES MELITUS**

LAPORAN TUGAS AKHIR



REFFY MAULANA
41521010103

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**IMPLEMENTASI *PARTICLE SWARM OPTIMIZATION* PADA
ALGORITMA *NAIVE BAYES*, *DECISION TREE C4.5*, DAN *RANDOM
FOREST* UNTUK KLASIFIKASI DIABETES MELITUS**

LAPORAN TUGAS AKHIR

REFFY MAULANA

41521010103

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Reffy Maulana

NIM : 41521010103

Program Studi : Teknik Informatika

Judul Laporan Skripsi : Implementasi *Particle Swarm Optimization* Pada Algoritma *Naive Bayes*, *Decision Tree C4.5* dan *Random Forest* Untuk Klasifikasi Diabetes Melitus

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Juli 2025



Reffy Maulana

HALAMAN PENGESAHAN

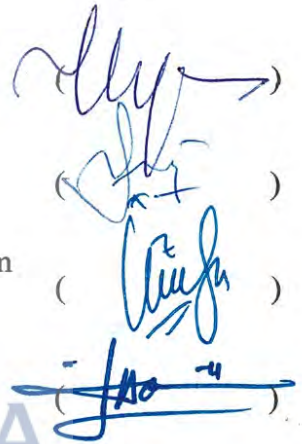
Laporan Skripsi ini diajukan oleh:

Nama : Reffy Maulana
NIM : 41521010103
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi *Particle Swarm Optimization* Pada Algoritma *Naive Bayes*, *Decision Tree C4.5* dan *Random Forest* Untuk Klasifikasi Diabetes Melitus

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Dr. Ir. Eliyani
NIDN : 0321026901
Ketua Penguji : Dr. Afiyati, S.Si., M.T.
NIDN : 0316106908
Penguji 1 : Inna Sabily Karima, S.Kom, M.Kom
NIDN : 0324018902
Penguji 2 : Lukman Hakim, ST, M.Kom
NIDN : 0327107701



Jakarta, 19 Juli 2025

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Dr. Ir. Eliyani selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Semua teman, baik diluar perkuliahan maupun seperjuangan kuliah yang selalu berbagi ilmu dan informasi serta memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 19 Juli 2025

Reffy Maulana

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Reffy Maulana
NIM : 415231010103
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi *Particle Swarm Optimization* Pada Algoritma *Naive Bayes*, *Decision Tree C4.5* dan *Random Forest* Untuk Klasifikasi Diabetes Melitus

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 19 Juli 2025
Yang menyatakan,



Reffy Maulana.

ABSTRAK

Nama : Reffy Maulana
NIM : 41521010103
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementasi *Particle Swarm Optimization* Pada Algoritma *Naive Bayes*, *Decision Tree C4.5* dan *Random Forest* Untuk Klasifikasi Diabetes Melitus
Dosen Pembimbing : Dr. Ir. Eliyani

Kadar gula darah tinggi merupakan gejala Diabetes Melitus, kondisi metabolik kronis yang disebabkan oleh kelainan dalam sintesis atau penggunaan insulin. Organisasi Kesehatan Dunia (WHO) memperkirakan bahwa 1,5 juta orang meninggal karena diabetes setiap tahun, dan lebih dari 422 juta orang di seluruh dunia menderita kondisi tersebut. Teknik yang andal untuk deteksi dini dan kategorisasi sangat penting untuk mendukung peningkatan manajemen penyakit karena peningkatan jumlah pasien diabetes telah sangat meningkatkan beban kesehatan global. Menggunakan data dari survei tahunan BRFSS 2023, studi ini membandingkan dan menilai seberapa baik kinerja algoritma *Naive Bayes*, *Decision Tree C4.5*, dan *Random Forest* dalam klasifikasi diabetes melitus. Metrik termasuk *accuracy*, *precision*, *recall*, dan *F1-score* digunakan dalam evaluasi untuk mengukur kinerja setiap algoritma. Pekerjaan ini juga menggabungkan teknik optimasi *Particle Swarm Optimization (PSO)* untuk meningkatkan kinerja klasifikasi dan mengurangi kesalahan. Hasil penelitian menunjukkan bahwa *Random Forest* tanpa optimasi PSO mencapai akurasi tertinggi sebesar 95.21%. Meskipun PSO berhasil meningkatkan akurasi *Naive Bayes* dari 80.74% menjadi 82.08% dan *C4.5* dari 91.24% menjadi 91.57%, *Random Forest* mengalami sedikit penurunan akurasi setelah dioptimasi menjadi 94.60%, mengindikasikan bahwa model tersebut sudah sangat optimal. Analisis fitur penting (*feature importance*) lebih lanjut mengungkapkan bahwa atribut seperti *GenHealth*, *BMI*, dan *Age* secara konsisten menjadi prediktor paling signifikan untuk klasifikasi Diabetes Melitus. Penelitian ini diharapkan dapat digunakan sebagai panduan untuk pembuatan aplikasi medis berbasis data yang akurat dan efektif, terutama yang membantu dalam deteksi dini diabetes. Hasilnya, karya ini memajukan penciptaan solusi teknologi berdasarkan *machine learning* untuk mengatasi berbagai masalah di bidang kesehatan global.

Kata kunci: Diabetes, Klasifikasi, *Naive Bayes*, *Decision Tree C4.5*, *Random Forest*, *Particle Swarm Optimization*

ABSTRACT

Nama : Reffy Maulana
NIM : 41521010103
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Implementation of Particle Swarm Optimization in Naive Bayes, Decision Tree C4.5, and Random Forest Algorithms for Diabetes Mellitus Classification
Dosen Pembimbing : Dr. Ir. Eliyani

High blood sugar levels are a symptom of diabetes mellitus, a chronic metabolic condition brought on by abnormalities in the synthesis or usage of insulin. The World Health Organization (WHO) estimates that 1.5 million people die from diabetes each year, and that over 422 million people worldwide have the condition. Reliable techniques for early detection and categorization are crucial to supporting improved disease management because the rise in diabetes patients has greatly increased the global health burden. Using data from the 2023 BRFSS annual survey, this study compares and assesses how well the Naive Bayes, Decision Tree C4.5, and Random Forest algorithms perform in the classification of diabetes mellitus. Metrics including accuracy, precision, recall, and F1-score are used in the evaluation to gauge each algorithm's performance. This work also incorporates an optimization technique based on Particle Swarm Optimization (PSO) to enhance classification performance and reduce error. The research results show that Random Forest without PSO optimization achieved the highest accuracy of 95.21%. Although PSO successfully improved Naïve Bayes accuracy from 80.74% to 82.08% and C4.5 from 91.24% to 91.57%, Random Forest experienced a slight decrease in accuracy after optimization to 94.60%, indicating that the model was already highly optimal. Further analysis of feature importance revealed that attributes such as GenHealth, BMI, and Age were consistently the most significant predictors for Diabetes Mellitus classification. It is anticipated that this study will be used as a guide for the creation of precise and effective data-driven medical applications, especially those that aid in the early detection of diabetes. As a result, this work advances the creation of technology solutions based on machine learning to address issues in the field of global health.

Kata kunci: Diabetes, Classification, Naive Bayes, Decision Tree C4.5, Random Forest, Particle Swarm Optimization

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terdahulu	9
2.2 Gap Penelitian.....	79
2.3 Teori Pendukung.....	80
2.3.1 Diabetes Melitus.....	81
2.3.2 <i>Machine Learning</i>	81
2.3.3 <i>Data Mining</i>	82
2.3.4 BRFSS 2023 <i>Dataset</i>	82
2.3.5 <i>Naïve Bayes</i>	82
2.3.6 <i>Decision Tree C4.5</i>	83
2.3.7 <i>Random Forest</i>	85
2.3.8 <i>Particle Swarm Optimization</i>	87
2.3.9 <i>Confusion Matrix</i>	88

2.3.10	<i>Correlation Matrix</i>	89
2.3.11	<i>Feature Importance</i>	90
2.3.12	<i>Permutation Importance</i>	90
BAB III METODE PENELITIAN		91
3.1	Jenis Penelitian	91
3.2	Tahapan Penelitian	92
3.2.1	<i>Data Collection</i>	92
3.2.2	<i>Pre-Processing Data</i>	93
3.2.3	<i>Model Implementation</i>	95
3.2.4	<i>Model Evaluation</i>	95
BAB IV PEMBAHASAN		97
4.1.	<i>Data Collection</i>	97
4.2.	<i>Pre-Processing Data</i>	97
4.2.1.	<i>Data Exploration</i>	97
4.2.2.	<i>Feature Selection</i>	98
4.2.3.	<i>Data Cleansing</i>	102
4.2.4.	<i>Data Transformation</i>	103
4.2.5.	<i>Handling Imbalance Data</i>	104
4.3.	<i>Model Implementation</i>	105
4.3.1.	<i>Naïve Bayes Implementation</i>	107
4.3.2.	<i>Decision Tree C4.5 Implementation</i>	109
4.3.3.	<i>Random Forest Implementation</i>	113
4.4.	<i>Model Evaluation</i>	116
4.4.1.	<i>Naïve Bayes Tanpa Optimasi</i>	117
4.4.2.	<i>Naïve Bayes Dengan Optimasi PSO</i>	120
4.4.3.	<i>Random Forest Tanpa Optimasi</i>	123
4.4.4.	<i>Random Forest Dengan Optimasi PSO</i>	126
4.4.5.	<i>Decision Tree C4.5 Tanpa Optimasi</i>	131
4.4.6.	<i>Decision Tree C4.5 Dengan Optimasi PSO</i>	133
4.4.7.	<i>Perbandingan Evaluasi Keseluruhan Model</i>	137
4.4.8.	<i>Analisis Komparatif Pentingnya Fitur</i>	139
4.4.9.	<i>Interpretasi Hasil Evaluasi Model</i>	141

BAB V KESIMPULAN DAN SARAN.....	143
5.1 Kesimpulan	143
5.2 Saran	144
DAFTAR PUSTAKA.....	146
LAMPIRAN.....	152



DAFTAR TABEL

Tabel 2.1 Tabel representasi confusion matrix	89
Tabel 4.1 Gambaran Dataset BRFSS 2023.....	98
Tabel 4.2 Pemilihan Atribut.....	98
Tabel 4.3 Distribusi data label Diabetes	104
Tabel 4.4 Distribusi data label Diabetes setelah SMOTENN.....	105
Tabel 4.5 Skema Pengujian.....	106
Tabel 4.6. Classification Report Naïve Bayes Tanpa PSO	118
Tabel 4.7. Classification Report Naïve Bayes Dengan PSO	121
Tabel 4.8. Hyperparameter Tuning Naïve Bayes Oleh PSO	121
Tabel 4.9. Classification Report Random Forest Tanpa PSO.....	124
Tabel 4.10. Classification Report Random Forest Dengan PSO	127
Tabel 4.11. Hyperparameter Tuning Random Forest Oleh PSO	128
Tabel 4.12. Classification Decision Tree C4.5 Tanpa PSO	132
Tabel 4.13. Classification Report Decision Tree C4.5 dengan PSO	134
Tabel 4.14. Hyperparameter Tuning Decision Tree C4.5 Oleh PSO.....	135
Tabel 4.15. Perbandingan Accuracy Keseluruhan Model	138
Tabel 4.16. Perbandingan Precision, Recall dan F1-Score	139



DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian	92
Gambar 4.1 Proses konversi dataset	97
Gambar 4.2 Heatmap Correlation Matrix antar fitur dan label.....	100
Gambar 4.3 Bar Chart Correlation Matrix antar fitur dan label.....	101
Gambar 4.4 Penghapusan nilai yang hilang dan duplikasi data.....	102
Gambar 4.5 Penghapusan nilai yang tidak valid pada label Diabetes	103
Gambar 4.6 Splitting Data dan Standarisasi Data.....	104
Gambar 4.7 Oversampling menggunakan SMOTENN	105
Gambar 4.8. Split Data Train dan Data Test.....	106
Gambar 4.9. Pembentukan Model Naïve Bayes	107
Gambar 4.10. Fungsi Objektif untuk PSO pada Naïve Bayes	108
Gambar 4.11. Eksekusi PSO dan Hasil Tuning Hyperparameter Naïve Bayes..	109
Gambar 4.12. Pembentukan Model Naïve Bayes dengan Hyperparameter Terbaik Hasil PSO.....	109
Gambar 4.13. Pembentukan Model Decision Tree C4.5	110
Gambar 4.14. Fungsi Objektif untuk PSO pada Decision Tree C4.5	111
Gambar 4.15. Eksekusi PSO dan Hasil Tuning Hyperparameter Decision Tree C4.5	112
Gambar 4.16. Pembentukan Model Decision Tree C4.5 dengan Hyperparameter Terbaik Hasil PSO	113
Gambar 4.17. Pembentukan Model Random Forest.....	113
Gambar 4.18. Fungsi Objektif untuk PSO pada Random Forest	114
Gambar 4.19. Eksekusi PSO dan Hasil Tuning Hyperparameter Random Forest	115
Gambar 4.20. Pembentukan Model Random Forest dengan Hyperparameter Terbaik Hasil PSO	116
Gambar 4.21. Confusion Matrix – Nave Bayes	117
Gambar 4.22. Urutan Pentingnya Fitur – Naïve Bayes Tanpa PSO	119
Gambar 4.23. Confusion Matrix – Naïve Bayes + PSO	120
Gambar 4.24. Urutan Pentingnya Fitur – Naïve Bayes + PSO.....	122
Gambar 4.25. Confusion Matrix – Random Forest.....	124
Gambar 4.26. Urutan Pentingnya Fitur - Random Forest Tanpa PSO.....	125
Gambar 4.27. Confusion Matrix – Random Forest + PSO.....	126
Gambar 4.28. Urutan Pentingnya Fitur – Random Forest + PSO.....	130
Gambar 4.29. Confusion Matrix – Decision Tree C4.5.....	131
Gambar 4.30. Urutan Pentingnya Fitur – Decision Tree C4.5 Tanpa PSO	133
Gambar 4.31. Confusion Matrix – Decision Tree C4.5 + PSO	134
Gambar 4.32. Urutan Pentingnya Fitur – Decision Tree C4.5 + PSO	137

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi	152
Lampiran 2. Curriculum Vitae	153
Lampiran 3. Surat Pernyataan HAKI	154
Lampiran 4. Sertifikat BNSP	156
Lampiran 5. Form Revisi Dosen Penguji	157
Lampiran 6. Hasil Cek Turnitin	159
Lampiran 7. Halaman Pernyataan Persetujuan Publikasi Tugas Akhir	160

