



**PENINGKATAN AKURASI PREDIKSI DIABETES MENGGUNAKAN RANDOM  
FOREST DAN XGBOOST DENGAN SELEKSI FITUR BERBASIS PSO DAN GA**

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**DZIRA NAUFIA JAWZA**

**NIM 2111016320009**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**MEI 2025**



**PENINGKATAN AKURASI PREDIKSI DIABETES MENGGUNAKAN RANDOM  
FOREST DAN XGBOOST DENGAN SELEKSI FITUR BERBASIS PSO DAN GA**

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**DZIRA NAUFIA JAWZA**

**NIM 2111016320009**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**MEI 2025**

# SKRIPSI

## PENINGKATAN AKURASI PREDIKSI DIABETES MENGGUNAKAN RANDOM FOREST DAN XGBOOST DENGAN SELEKSI FITUR BERBASIS PSO DAN GA

Oleh:

**DZIRA NAUFIA JAWZA**

**NIM. 2111016320009**

Telah dipertahankan di depan Dosen Penguji pada tanggal 08 Mei 2025.

Susunan Dosen Penguji:

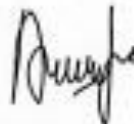
**Pembimbing I**



Muhammad Itqan Mazdadi, S.Kom., M.Kom.

NIP. 199006122019031013

**Dosen Penguji I**



Triando Hamonangan Saragih, S.Kom., M.Kom.

NIP. 199308242019031012

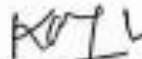
**Pembimbing II**



Andi Farmadi, S.Si., M.T.

NIP. 197307252008011006

**Dosen Penguji II**



Dwi Kartini, S.Kom., M.Kom.

NIP. 1987042120121220033

Banjarbaru, 09 Mei 2025

**Koordinator Program Studi Ilmu Komputer**



Dwi Kartini, S.Kom., M.Kom.

NIP. 1987042120121220033

## SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam jurnal ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 8 Mei 2025

Yang Menyatakan,



Dzira Naufia Jawza

NIM. 2111016320009

## ABSTRAK

### **PENINGKATAN AKURASI PREDIKSI DIABETES MENGGUNAKAN RANDOM FOREST DAN XGBOOST DENGAN SELEKSI FITUR BERBASIS PSO DAN GA**

(Oleh : Dzira Naufia Jawza; Pembimbing: Muhammad Itqan Mazdadi, S.Kom., M.Kom. dan Andi Farmadi, S.Si., M.T.; 2025; halaman)

Diabetes adalah tantangan kesehatan global yang kritis, diklasifikasikan sebagai penyakit tidak menular, yang memengaruhi lebih dari 422 juta orang di seluruh dunia, dengan tingkat prevalensi terus meningkat setiap tahunnya. Studi ini membahas kebutuhan untuk prediksi diabetes yang lebih akurat dengan mengevaluasi kinerja algoritma klasifikasi Random Forest dan Extreme Gradient Boosting (XGBoost) pada kumpulan data diabetes yang tersedia untuk umum dari Kaggle. Penelitian ini bertujuan untuk meningkatkan akurasi prediksi melalui teknik pemilihan fitur. Kumpulan data tersebut terdiri dari 768 catatan dengan 9 atribut, termasuk indikator medis seperti kehamilan, kadar glukosa, tekanan darah, dan BMI, dengan label target yang mengkategorikan hasil sebagai diabetes (1) atau non-diabetes (0). Prapemrosesan dilakukan untuk menangani data yang hilang, memastikan keandalan data. Metode pemilihan fitur, yaitu Particle Swarm Optimization (PSO) dan Genetic Algorithm (GA), digunakan untuk mengidentifikasi atribut yang paling relevan, meningkatkan efisiensi dan akurasi model. Temuan tersebut mengungkapkan bahwa tanpa pemilihan fitur, model Random Forest mencapai skor Area Under Curve (AUC) sebesar 0,8120, sementara XGBoost memperoleh skor 0,7666. Setelah menerapkan pemilihan fitur berbasis PSO, skor AUC meningkat menjadi 0,8582 dan 0,8250 untuk Random Forest dan XGBoost. Pemilihan fitur berbasis GA selanjutnya meningkatkan skor ini menjadi 0,8612 untuk Random Forest dan 0,8351 untuk XGBoost, yang menunjukkan peningkatan hingga 8,9%. Hasil ini menyoroti efektivitas GA dalam mengungguli PSO untuk pemilihan fitur. Studi ini menggarisbawahi pentingnya mengintegrasikan teknik pemilihan fitur dalam meningkatkan akurasi model klasifikasi. Temuan ini memiliki implikasi praktis untuk mengembangkan alat prediksi yang kuat untuk deteksi diabetes dini, yang dapat memfasilitasi diagnosis yang tepat waktu dan akurat dalam pengaturan klinis.

**Kata kunci:** Random Forest, XGBoost, PSO, GA, Diabetes

## **ABSTRACT**

### **ENHANCING DIABETES PREDICTION ACCURACY USING RANDOM FOREST AND XGBOOST WITH PSO AND GA-BASED FEATURE SELECTION**

(By : Dzira Naufia Jawza; Supervisor: Muhammad Itqan Mazdadi, S.Kom., M.Kom. and Andi Farmadi, S.Si., M.T.; 2025; pages)

*Diabetes is a critical global health challenge, classified as a non-communicable disease, affecting over 422 million individuals worldwide, with prevalence rates continuing to rise annually. This study addresses the need for more accurate diabetes prediction by evaluating the performance of Random Forest and Extreme Gradient Boosting (XGBoost) classification algorithms on a publicly available diabetes dataset from Kaggle. The research aims to improve prediction accuracy through feature selection techniques. The dataset comprises 768 records with 9 attributes, including medical indicators such as pregnancies, glucose levels, blood pressure, and BMI, with the target label categorizing outcomes as diabetic (1) or non-diabetic (0). Preprocessing was conducted to handle missing data, ensuring data reliability. Feature selection methods, namely Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), were employed to identify the most relevant attributes, enhancing the models' efficiency and accuracy. The findings revealed that without feature selection, the Random Forest model achieved an Area Under Curve (AUC) score of 0.8120, while XGBoost scored 0.7666. After applying PSO-based feature selection, the AUC scores increased to 0.8582 and 0.8250 for Random Forest and XGBoost, respectively. GA-based feature selection further improved these scores to 0.8612 for Random Forest and 0.8351 for XGBoost, demonstrating an improvement of up to 8.9%. These results highlight the effectiveness of GA in outperforming PSO for feature selection. This study underscores the significance of integrating feature selection techniques in enhancing classification model accuracy. The findings hold practical implications for developing robust predictive tools for early diabetes detection, which can facilitate timely and precise diagnoses in clinical settings.*

**Keywords:** Random Forest, XGBoost, PSO, GA, Diabetes

## KATA PENGANTAR

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan jurnal yang berjudul “*The Enhancing Diabetes Prediction Accuracy Using Random Forest and XGBoost with PSO and GA-Based Feature Selection*” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada semua pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan jurnal ini, adapun yang dimaksud adalah sebagai berikut:

1. Allah SWT yang telah memberikan kesehatan, kekuatan, dan kemudahan dalam menyelesaikan jurnal ini.
2. Diri sendiri atas tekad dan semangat yang tak pernah pudar meskipun menghadapi berbagai tantangan dan rintangan selama proses penyusunan jurnal ini.
3. Keluarga tercinta yang senantiasa menjadi sumber motivasi terbesar, memberikan doa, dukungan dalam penyelesaian jurnal ini dan kasih sayang tanpa henti.
4. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku dosen pembimbing utama yang telah memberikan arahan, bimbingan, serta waktunya dalam membantu kelancaran penyusunan jurnal ini.
5. Bapak Andi Farmadi, S.Si., M.T. selaku dosen pembimbing pendamping yang dengan penuh kesabaran turut membimbing dan memberikan masukan demi kesempurnaan jurnal ini.
6. Ibu Dwi Kartini, S.Kom., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau jurnal ini dapat diselesaikan.
7. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
8. Teman-teman dan sahabat-sahabat Ilmu Komputer yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan jurnal.
9. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah berkontribusi dalam membantu penyelesaian jurnal ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 8 Mei 2025

A handwritten signature in black ink, appearing to read 'Dzira Naufia Jawza', written in a cursive style.

Dzira Naufia Jawza