



**EFFECT OF DATASET QUANTITY AND FEATURE VARIABILITY  
ON MACHINE LEARNING PERFORMANCE FOR ECG-BASED  
DIABETES PREDICTION**

**Skripsi**

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

**Oleh**

**MUHAMAD RAMADHANI**

**NIM 2111016310013**

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

**SEPTEMBER 2025**



**EFFECT OF DATASET QUANTITY AND FEATURE VARIABILITY ON  
MACHINE LEARNING PERFORMANCE FOR ECG-BASED DIABETES  
PREDICTION**

**Skripsi**

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

**Oleh**

**MUHAMAD RAMADHANI**

**NIM 2111016310013**

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

**SEPTEMBER 2025**

# SKRIPSI

## EFFECT OF DATASET QUANTITY AND FEATURE VARIABILITY ON MACHINE LEARNING PERFORMANCE FOR ECG-BASED DIABETES PREDECTION

Oleh:

**MUHAMAD RAMADHANI**

**NIM. 2111016310013**

Telah dipertahankan di depan Dosen Penguji pada tanggal 17 September 2025.

Susunan Dosen Penguji:

**Pembimbing I**

**Dosen Penguji I**



Mohammad Reza Faisal, S.T., M.T., Ph.D

NIP. 197612202008121001



Muliadi, S.Kom, M.Sc

NIP. 197804222010121002

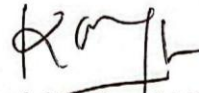
**Pembimbing II**

**Dosen Penguji II**



Dodon Turianto Nugrahadi, S.Kom, M.Eng

NIP. 198001122009121002



Dwi Kartini, S.Kom, M.Kom

NIP. 1987042120121220033

Banjarbaru, 18 September 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, 18 September 2025

Yang Menyatakan,



Muhamad Ramadhani  
NIM. 2111016310013

## ABSTRAK

### **PENGARUH JUMLAH DATASET DAN VARIASI FITUR TERHADAP KINERJA MACHINE LEARNING UNTUK PREDIKSI DIABETES BERBASIS ECG**

(Oleh: Muhamad Ramadhani; Pembimbing: Mohammad Reza Faisal, S.T, M.T, Ph.D., Dodon Turianto Nugrahadi, S.Kom, M.Eng.; 2025; halaman)

Diabetes adalah penyakit kronis yang sering kali tidak terdeteksi pada tahap awal, sehingga berpotensi menimbulkan komplikasi serius apabila tidak ditangani. Metode skrining awal yang non-invasif dan andal diperlukan untuk meningkatkan kepatuhan pasien dibandingkan teknik pemantauan invasif tradisional. Meskipun sinyal ECG telah banyak diteliti untuk deteksi penyakit jantung, penerapannya dalam klasifikasi diabetes masih kurang dieksplorasi, dengan studi sebelumnya pada dataset DINAMO melaporkan akurasi yang terbatas. Penelitian ini mengatasi kesenjangan tersebut dengan menerapkan deteksi R-peak untuk segmentasi ECG dan secara sistematis mengevaluasi pengaruh jumlah data dan variasi fitur terhadap kinerja model. Tiga algoritma machine learning klasik (SVM, KNN, dan DT) dievaluasi pada berbagai konfigurasi jumlah data dan variasi fitur. Hasil eksperimen menunjukkan bahwa DT mencapai akurasi tertinggi sebesar 85,30% dengan 5000 sampel dan 500 fitur ECG. Temuan ini menegaskan kemampuan algoritma berbasis pohon dalam memberikan prediksi diabetes yang presisi dan efisien secara non-invasif. Hasil penelitian ini menjadi dasar untuk eksplorasi lebih lanjut terhadap metode yang lebih maju guna meningkatkan ketahanan dan adaptabilitas dalam praktik klinis.

**Kata kunci:** prediksi diabetes, sinyal ECG, segmentasi R-peak, machine learning

## ABSTRACT

### **EFFECT OF DATASET QUANTITY AND FEATURE VARIABILITY ON MACHINE LEARNING PERFORMANCE FOR ECG-BASED DIABETES PREDICTION**

(By: Muhamad Ramadhani; Supervisors: Mohammad Reza Faisal, S.T, M.T, Ph.D., Dodon Turianto Nugrahadi, S.Kom, M.Eng.; 2025; pages)

Diabetes is a chronic disease that often goes undetected in its early stages, potentially leading to severe complications if left untreated. Non-invasive and reliable early screening methods are required to improve patient compliance compared to traditional invasive monitoring techniques. While ECG signals have been extensively studied for heart disease detection, their application for diabetes classification remains underexplored, with previous studies on the DINAMO dataset reporting limited accuracy. This study addresses this gap by applying R-peak detection for ECG segmentation and systematically evaluating the effect of dataset quantity and feature variability on model performance. Three classical machine learning algorithms (SVM, KNN, and DT) were evaluated under multiple dataset quantity and feature variability configurations. Experimental results indicate that DT achieved the highest accuracy of 85.30% with 5000 samples and 500 ECG features. These results underscore the capability of tree-based algorithms to deliver precise and efficient non-invasive diabetes predictions. The outcomes provide a foundation for further exploration of advanced methods to enhance robustness and adaptability in clinical practice.

**Keywords:** diabetes prediction, ECG signal, R-peak segmentation, machine learning

## 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 “*Effect of Dataset Quantity and Feature Variability on Machine Learning Performance for ECG-Based Diabetes Prediction*” 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. Kepada diri sendiri atas tekad dan semangat yang terus terjaga meskipun menghadapi berbagai tantangan dan hambatan dalam penyusunan jurnal ini.
3. Keluarga tercinta Mama dan Abah, yang selalu memberikan doa, dukungan, dan kasih sayang.
4. Keluarga besar yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian jurnal ini.
5. Bapak Mohammad Reza Faisal, S.Si, S.T, M.T, Ph.D. selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
6. Bapak Dodon Turianto Nugrahadi, S.Kom, M.Eng. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
7. Ibu Dwi Kartini, S.Kom., M.Kom. selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau jurnal ini dapat diselesaikan.
8. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
9. Muhammad Zaisya Fitriannuur Rahman, selaku partner penelitian yang telah memberikan dukungan, kerja sama, serta kontribusi berharga dalam penyusunan penelitian ini.

10. Teman-teman Ryzen<sup>21</sup> serta sahabat-sahabat keluarga Ilmu Komputer yang telah memberikan dukungan, bantuan, dan doa selama proses penyusunan jurnal ini.
11. 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 bermanfaat bagi pengembangan ilmu pengetahuan dan memberikan manfaat bagi para pembaca, serta mendapat keridhaan Allah SWT.

Banjarbaru, 18 September 2025



Muhamad Ramadhani

