



**KLASIFIKASI ARITMIA DARI DATA ELEKTROKARDIOGRAM  
MENGUNAKAN *RANDOM FOREST* DENGAN SELEKSI FITUR  
*INFORMATION GAIN***

**Skripsi**

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

**Oleh**

**SITI BADRIYAH**

**NIM 1611016120014**

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

**JULI 2023**



**KLASIFIKASI ARITMIA DARI DATA ELEKTROKARDIOGRAM  
MENGUNAKAN *RANDOM FOREST* DENGAN SELEKSI FITUR  
*INFORMATION GAIN***

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Sarjana Strata – 1 Ilmu Komputer**

**Oleh**

**SITI BADRIYAH**

**NIM 1611016120014**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
JULI 2023**

# SKRIPSI

## KLASIFIKASI ARITMIA DARI DATA ELEKTROKARDIOGRAM MENGUNAKAN *RANDOM FOREST* DENGAN SELEKSI FITUR *INFORMATION GAIN*

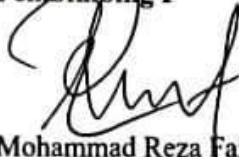
Oleh :

**SITI BADRIYAH**  
**NIM 1611016120014**

Telah dipertahankan di depan Dosen Penguji pada tanggal 20 Juli 2023.

Susunan Dosen Penguji :

**Pembimbing I**



Mohammad Reza Faisal, S.T., M.T., Ph.D  
NIP. 197612202008121001

**Dosen Penguji I**



Irwan Budiman, S.T., M.Kom  
NIP. 197703252008121001

**Pembimbing II**



Dwi Kartini, S.Kom., M.Kom  
NIP. 198704212012122003

**Dosen Penguji II**



Muhammad Itqan Mazdadi, S.Kom., M.Kom  
NIP. 199006122019031013



Banjarmasin, Juli 2023  
Ketua Program Studi Ilmu Komputer

Irwan Budiman, S.T., M.Kom  
197703252008121001

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa dalam skripsi 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, Juli 2023



**Siti Badriyah**

**NIM. 1611016120014**

## ABSTRAK

**KLASIFIKASI ARITMIA DARI DATA ELEKTROKARDIOGRAM MENGGUNAKAN *RANDOM FOREST* DENGAN SELEKSI FITUR *INFORMATION GAIN*** (Oleh: Siti Badriyah; Pembimbing: Mohammad Reza Faisal, S. T., M.T., Ph.D dan Dwi Kartini, S.Kom., M.kom; 40 Halaman)

Aritmia adalah kondisi jantung yang berdetak tidak normal. Jantung dapat berdetak lebih cepat, lebih lambat atau tidak teratur. Kondisi tidak normal ini menyebabkan jantung tidak dapat memompa darah dengan efektif dan dapat mengganggu organ-organ tubuh lain. Salah satu cara untuk mendeteksi lebih dini penyakit aritmia ini adalah dengan menggunakan machine learning. Penelitian ini menggunakan random forest untuk mengenali pola dari sebuah data pelatihan arrhythmia. Data pelatihan terbagi menjadi 4 kelas yaitu kelas normal, kelas Atrial Fibrillation, kelas PVC Bigeminy dan kelas ventricular tachycardia dengan masing-masing kelas memiliki 2160 fitur. Karena banyaknya fitur yang dimiliki data sehingga bisa membuat pelatihan machine learning kurang maksimal sehingga diperlukannya sebuah metode seleksi fitur yang salah satu metodenya adalah information gain. Berdasarkan hasil penelitian, seleksi fitur information dapat meningkatkan akurasi klasifikasi random forest sebesar 3,57 % dari 60,71 % yang tanpa menggunakan seleksi fitur menjadi sebesar 64,28 % yang dengan menggunakan seleksi fitur information gain.

**Kata Kunci :** Arrhythmia Dataset, Random Forest, Information Gain, Seleksi Fitur

## ABSTRACT

***Classification of Arrhythmia from Electrocardiogram Data Using Random Forest with Information Gain Feature Selection*** (by: Siti Badriyah; supervisors: Mohammad Reza Faisal, S. T., M.T., Ph.D and Dwi Kartini, S.Kom., M.kom; 40 pages)

*Arrhythmia is a condition where the heart beats abnormally. The heart can beat faster, slower, or irregularly. This abnormal condition prevents the heart from effectively pumping blood and can disrupt other organs in the body. One way to detect this arrhythmia disease early is by using machine learning. This research uses random forest to recognize patterns from a training dataset of arrhythmia. The training data is divided into 4 classes: normal class, Atrial Fibrillation class, PVC Bigeminy class, and ventricular tachycardia class, each with 2160 features. Due to the high number of features in the data, it can hinder the effectiveness of machine learning training, thus requiring a feature selection method, one of which is information gain. Based on the research results, information gain feature selection can improve the classification accuracy of random forest by 3.57%, from 60.71% without using feature selection to 64.28% with feature selection information gain.*

***Keyword :*** Arrhythmia Dataset, Random Forest, Information Gain, Feature Selection

## PRAKATA

Puji dan syukur kepada Allah SWT atas berkat rahmat-Nya penulis dapat menyelesaikan skripsi yang berjudul **Klasifikasi Aritmia Dari Data Elektrokardiogram Menggunakan *Random Forest* Dengan Seleksi Fitur *Information Gain*** untuk memenuhi syarat dalam menyelesaikan pendidikan program Strata-1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Tidak lupa penulis menyampaikan ucapan terimakasih kepada pihak-pihak terkait yang sangat mendukung dan membantu penulis dalam penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut:

1. Allah SWT, karena atas limpahan karunia-Nya lah penulis dapat menyelesaikan skripsi ini.
2. Keluarga yang senantiasa memberikan doa, semangat, dukungan, hingga kepercayaan yang membuat penulis selalu bekerja keras menyelesaikan skripsi.
3. Bapak Mohammad Reza Faisal S.T., M.T. Ph.D selaku dosen pembimbing utama dan Ibu Dwi Kartini, S.Kom., M.Kom. selaku dosen pembimbing pendamping yang senantiasa membimbing, membantu, dan meluangkan waktu dalam proses penyelesaian skripsi ini.
4. Bapak Irwan Budiman, S.T., M.Kom. dan Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku dosen penguji yang telah memberikan masukan-masukan untuk penyajian skripsi.
5. Bapak Irwan Budiman, S.T., M.Kom. selaku ketua program studi Ilmu Komputer beserta seluruh dosen dan karyawan/staff pegawai Fakultas MIPA Universitas Lambung Mangkurat atas bantuan yang diberikan selama penulis mengikuti studi.
6. Teman-teman keluarga Ilmu Komputer angkatan 2016 serta sahabat penulis yang telah banyak memberikan bantuan serta motivasi sehingga skripsi ini dapat terselesaikan.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam menyelesaikan skripsi ini.

Akhir kata penulis menyadari bahwa skripsi ini jauh dari sempurna, kritik dan saran yang membangun sangat diperlukan. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Juli 2023

A handwritten signature in black ink, appearing to read 'Siti Badriyah', with a large, stylized flourish at the end.

Siti Badriyah



## DAFTAR ISI

PERNYATAAN.....	ii
ABSTRAK.....	iv
ABSTRACT.....	v
PRAKATA.....	vi
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I 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
BAB II TINJAUAN PUSTAKA.....	4
2.1 Kajian Terdahulu.....	4
2.2 Keaslian Penelitian.....	6
2.3 Landasan Teori.....	7
2.3.1 Penyakit Jantung.....	7
2.3.2 Elektrokardiogram.....	7
2.3.3 Detak Jantung Normal.....	9
2.3.4 Aritmia.....	10
2.3.5 Atrial Fibrillation.....	10
2.3.6 PVC Bigeminy.....	11
2.3.7 Ventricular Tachycardia.....	12
2.3.8 Data Mining.....	12
2.3.9 Klasifikasi.....	14
2.3.10 Seleksi Fitur.....	15
2.3.11 Information Gain.....	17
2.3.12 Random Forest.....	18
2.3.13 Confusion Matrix.....	25
BAB III METODE PENELITIAN.....	27
3.1. Alat Penelitian.....	27

3.2.	Bahan Penelitian .....	27
3.3.	Variabel Penelitian.....	27
3.4.	Prosedur Penelitian .....	28
BAB IV HASIL DAN PEMBAHASAN .....		30
4.1	Hasil.....	30
4.1.1	Pengumpulan Data .....	30
4.1.2	Pembagian Data .....	31
4.1.3	Seleksi Fitur Information Gain.....	32
4.1.4	Klasifikasi .....	35
4.1.5	Evaluasi .....	36
4.2	Pembahasan .....	38
BAB V PENUTUP.....		40
5.1	Kesimpulan.....	40
5.2	Saran.....	40
DAFTAR PUSTAKA .....		41
LAMPIRAN.....		45

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
Tabel 1 Keaslian Penelitian.....	6
Tabel 2 Rancangan Penelitian.....	7
Tabel 3. Confusion Matrix .....	26
Tabel 4 Ringkasan Dataset.....	29
Tabel 5 Hasil Pengumpulan Data Arrhythmia .....	30
Tabel 6 Sampel Dataset Arrhythmia .....	30
Tabel 7 Pembagian Data Training dan testing .....	31
Tabel 8 Pembagian Data Training.....	31
Tabel 9 Pembagian Data Testing .....	32
Tabel 10 Fitur Arrhythmia yang Memiliki Bobot .....	32
Tabel 11 Perangkingan Bobot Information Gain .....	33
Tabel 12 Pembagian Data Training Dengan Seleksi Fitur.....	34
Tabel 13 Pembagian Data Testing Dengan Seleksi Fitur.....	34
Tabel 14 Perbandingan Pemilihan Jumlah Seleksi Fitur Pada Klasifikasi .....	35
Tabel 15 Confusion Matrix Klasifikasi Random Forest Tanpa Seleksi Fitur .....	37
Tabel 16 Evaluasi Akurasi Random Forest Tanpa Seleksi Fitur .....	37
Tabel 17 Confusion Matrix Random Forest Dengan Seleksi Fitur 1800.....	37
Tabel 18 Evaluasi Akurasi Random Forest Dengan Seleksi Fitur 1800 .....	38

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
Gambar 1. Kompleks ekg detak jantung normal.....	8
Gambar 2. Contoh EKG dengan detak jantung normal .....	9
Gambar 3. EKG Atrial Fibrillation .....	11
Gambar 4. EKG PVC bigeminy.....	11
Gambar 5. EKG Ventricular Tachycardia.....	12
Gambar 6. Tahapan Data Mining.....	13
Gambar 7. Contoh Random Forest.....	21
Gambar 8. Random Forest .....	23
Gambar 9. Alur Penelitian.....	28
Gambar 10 Perbandingan Pemilihan Jumlah Seleksi Fitur Pada Klasifikasi.....	36
Gambar 11 Perbandingan Evaluasi Tanpa Seleksi Fitur dan Seleksi Fitur.....	39

## DAFTAR LAMPIRAN

<b>Lampiran</b>	<b>Halaman</b>
Lampiran 1 Source Code Klasifikasi Random Forest Tanpa Seleksi Fitur .....	46
Lampiran 2 Source Code Klasifikasi Random Forest Dengan Seleksi Fitur Information Gain.....	47