

**DISERTASI**

**OPTIMASI PEMILIHAN FITUR BERBASIS ALGORITMA  
METAHEURISTIK DAN *MACHINE LEARNING***



**ANNISA DARMAWAHYUNI**

**NIM: 03013622328005**

**BKU TEKNIK INFORMATIKA**

**PROGRAM STUDI DOKTOR ILMU TEKNIK  
PROGRAM PASCA SARJANA FAKULTAS TEKNIK  
UNIVERSITAS SRIWIJAYA**

**2025**

**HALAMAN PENGESAHAN**  
**OPTIMASI PEMILIHAN FITUR BERBASIS ALGORITMA**  
**METAHEURISTIK DAN *MACHINE LEARNING***

**DISERTASI**

Diajukan Untuk Memenuhi Salah Satu Syarat Memperoleh Gelar  
Doktor Ilmu Teknik Pada Program Pascasarjana  
Fakultas Teknik Universitas Sriwijaya

Disusun Oleh:

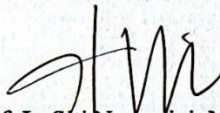
**ANNISA DARMAWAHYUNI**

**NIM: 03013622328005**

Telah Disetujui:

Palembang, Februari 2025

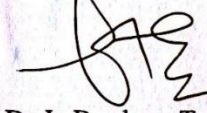
Promotor



Prof. Ir. Siti Nurmaini, M.T., Ph.D

NIP. 196908021994012001

Ko-Promotor



Prof. Dr. Ir. Bambang Tutuko, M.T

NIP. 196001121989031002

Mengetahui:

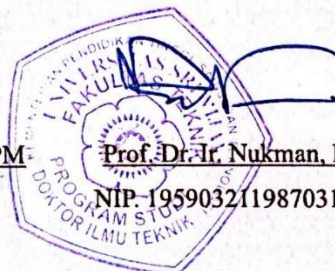
Dekan Fakultas Teknik



Dr. Ir. Bhakti Yudho Suprpto, ST, MT, IPM

NIP. 197502112003121002

Koordinator Program Studi



Prof. Dr. Ir. Nukman, MT

NIP. 195903211987031001



## HALAMAN PERSETUJUAN

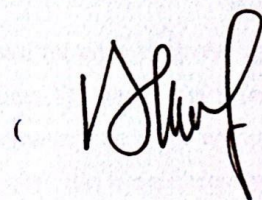
Dengan ini menyatakan bahwa disertasi Annisa Darmawahyuni yang berjudul "OPTIMASI PEMILIHAN FITUR BERBASIS ALGORITMA METAHEURISTIK DAN *MACHINE LEARNING*" telah dipertahankan di hadapan Sidang Ujian Tertutup Program Studi Ilmu Teknik Program Doktor, Fakultas Teknik, Universitas Sriwijaya pada Kamis, 30 Januari 2025.

Palembang, Januari 2025

Ditandatangani oleh Tim Penguji,

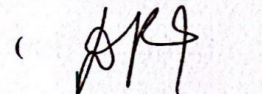
Ketua Tim Penguji:

Dr. Ir. Bhakti Yudho Suprpto, ST., MT., IPM  
NIP. 197502112003121002

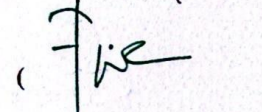
(  )

Anggota Tim Penguji:

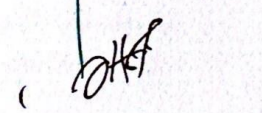
1. Dian Palupi Rini, M.Kom., Ph.D  
NIP. 197802232006042002

(  )

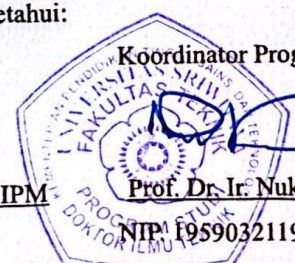
2. Dr. Firdaus, S.T., M.Kom  
NIP. 197801212008121003

(  )

3. Dr. Yesi Novaria Kunang, S.T., M.Kom.  
NIP.

(  )

  
Mengetahui:  
Dekan Fakultas Teknik  
Dr. Ir. Bhakti Yudho Suprpto, ST., MT., IPM  
NIP. 197502112003121002

  
Koordinator Program Studi  
Prof. Dr. Ir. Nukman, MT  
NIP. 195903211987031001



## PERNYATAAN ORISINALITAS/PLAGIARISME

Nama : Annisa Darmawahyuni  
NIM : 03013622328005  
Judul : Optimasi Pemilihan Fitur Berbasis Algoritma Metaheuristik dan  
*Machine Learning*

Dengan ini saya menyatakan keaslian disertasi ini. Disertasi ini dibimbing oleh satu Promotor dan Ko-promotor dan tidak melibatkan plagiarisme. Jika ditemukan adanya plagiarisme dalam disertasi ini, saya bersedia menerima sanksi akademik apa pun sesuai regulasi yang telah ditetapkan oleh Universitas Sriwijaya atas konsekuensinya.

Palembang, Februari 2025



Annisa Darmawahyuni

NIM. 03013622328005

## SURAT KETERANGAN PENGECEKAN SIMILARITY

Saya yang bertanda tangan di bawah ini:

Nama : Annisa Darmawahyuni

NIM : 03013622328005


Program Studi : Ilmu Teknik

Menyatakan bahwa benar hasil pengecekan similarity Disertasi yang berjudul **Optimasi Pemilihan Fitur Berbasis Algoritma Metaheuristik dan *Machine Learning*** adalah 5%.

Dicek oleh operator UPT Perpustakaan Universitas Sriwijaya melalui laman website <https://digilib.unsri.ac.id/>.

Demikian surat keterangan ini saya buat dengan sebenarnya dan dapat saya pertanggungjawabkan.

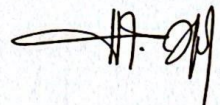
Mengetahui,



Prof. Ir. Siti Nurmaini, M.T., Ph.D  
NIP. 196908021994012001

Palembang, Februari 2025

Yang Menyatakan,



Annisa Darmawahyuni  
NIM. 03013622328005



**SURAT KETERANGAN  
PERBAIKAN LAPORAN PROPOSAL**

Dosen Pembahas Seminar Laporan Disertasi Program Studi Doktor Ilmu Teknik  
Fakultas Teknik Universitas Sriwijaya:

Nama : Dr. Ir. Bhakti Yudho Suprpto, ST., MT., IPM

NIP : 197502112003121002

menyatakan bahwa mahasiswa sebagai berikut:

Nama : Annisa Darmawahyuni

NIM : 03013622328005

BKU : Teknik Informatika

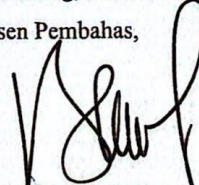
Judul Disertasi : OPTIMASI PEMILIHAN FITUR BERBASIS  
ALGORITMA METAHEURISTIK DAN *MACHINE  
LEARNING*

Telah melakukan perbaikan laporan disertasi sesuai dengan yang telah disarankan oleh  
Dosen Pembahas pada berita acara laporan disertasi.

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Palembang, Februari 2025

Dosen Pembahas,



Dr. Ir. Bhakti Yudho Suprpto, ST., MT., IPM

NIP. 197502112003121002

**SURAT KETERANGAN  
PERBAIKAN LAPORAN PROPOSAL**

Dosen Pembahas Seminar Laporan Disertasi Program Studi Doktor Ilmu Teknik  
Fakultas Teknik Universitas Sriwijaya:

Nama : Dian Palupi Rini, M.Kom., Ph.D

NIP : 197802232006042002

menyatakan bahwa mahasiswa sebagai berikut:

Nama : Annisa Darmawahyuni

NIM : 03013622328005

BKU : Teknik Informatika

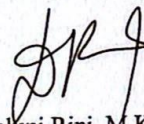
Judul Disertasi : OPTIMASI PEMILIHAN FITUR BERBASIS  
ALGORITMA METAHEURISTIK DAN *MACHINE  
LEARNING*

Telah melakukan perbaikan laporan disertasi sesuai dengan yang telah disarankan oleh  
Dosen Pembahas pada berita acara laporan disertasi.

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Palembang, Februari 2025

Dosen Pembahas,



Dian Palupi Rini, M.Kom., Ph.D

NIP. 197802232006042002



**SURAT KETERANGAN  
PERBAIKAN LAPORAN PROPOSAL**

Dosen Pembahas Seminar Laporan Disertasi Program Studi Doktor Ilmu Teknik  
Fakultas Teknik Universitas Sriwijaya:

Nama : Dr. Firdaus, S.T., M.Kom

NIP : 197801212008121003

menyatakan bahwa mahasiswa sebagai berikut:

Nama : Annisa Darmawahyuni

NIM : 03013622328005

BKU : Teknik Informatika

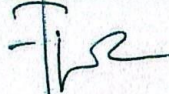
Judul Disertasi : OPTIMASI PEMILIHAN FITUR BERBASIS  
ALGORITMA METAHEURISTIK DAN *MACHINE  
LEARNING*

Telah melakukan perbaikan laporan disertasi sesuai dengan yang telah disarankan oleh  
Dosen Pembahas pada berita acara laporan disertasi.

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Palembang, Februari 2025

Dosen Pembahas,



Dr. Firdaus, S.T., M.Kom

NIP. 197801212008121003



**SURAT KETERANGAN  
PERBAIKAN LAPORAN PROPOSAL**

Dosen Pembahas Seminar Laporan Disertasi Program Studi Doktor Ilmu Teknik  
Fakultas Teknik Universitas Sriwijaya:

Nama : Dr. Yesi Novaria Kunang, S.T., M.Kom.

NIP : -

menyatakan bahwa mahasiswa sebagai berikut:

Nama : Annisa Darmawahyuni

NIM : 03013622328005

BKU : Teknik Informatika

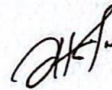
Judul Disertasi : OPTIMASI PEMILIHAN FITUR BERBASIS  
ALGORITMA METAHEURISTIK DAN *MACHINE  
LEARNING*

Telah melakukan perbaikan laporan disertasi sesuai dengan yang telah disarankan oleh  
Dosen Pembahas pada berita acara laporan disertasi.

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Palembang, Februari 2025

Dosen Pembahas,



Dr. Yesi Novaria Kunang, S.T., M.Kom.

NIP. -

## ABSTRAK

*Informasi klinis mencakup data medis dan demografi yang berperan dalam diagnosis. Seiring meningkatnya kompleksitas dan volume data medis, pemilihan fitur yang relevan menjadi tantangan utama dalam analisis klinis. Metode konvensional dalam pemilihan fitur membutuhkan waktu yang lama dan fitur suboptimal. Oleh karena itu, pendekatan otomatis berbasis komputer menjadi solusi yang lebih akurat.*

*Penelitian ini mengembangkan metodologi end-to-end berbasis algoritma metaheuristik dan supervised learning untuk optimasi pemilihan fitur pada data medis. Metodologi ini mencakup seluruh proses, mulai dari pemrosesan data mentah hingga evaluasi kinerja model. Pengujian pada penelitian ini menggunakan data terstruktur dan tidak terstruktur dari 11 dataset dengan 10 algoritma metaheuristik dan 4 algoritma supervised learning. Hasil pengujian dari 157 model menunjukkan bahwa pada data terstruktur, model mencapai kinerja di atas 90% untuk data berdimensi rendah dan 100% untuk data berdimensi sedang hingga tinggi pada semua metrik kinerja. Sementara itu, pada data tidak terstruktur, model mencapai akurasi 100% dalam klasifikasi biner dan multi kelas. Hasil penelitian ini membuktikan bahwa pemilihan fitur berbasis algoritma metaheuristik dapat meningkatkan kinerja model supervised learning pada data medis. Dengan demikian, metode yang dikembangkan diharapkan dapat mendukung pengambilan keputusan klinis yang lebih akurat dan andal.*

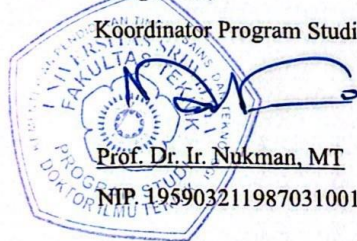
**Kata kunci:** algoritma metaheuristik, data medis, *fitness function*, supervised learning

Mengetahui,

Koordinator Program Studi

Prof. Dr. Ir. Nukman, MT

NIP. 195903211987031001



Promotor,

Prof. Ir. Siti Nurmaini, M.T., Ph.D

NIP. 196908021994012001



## ABSTRACT

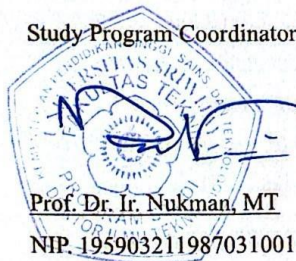
*Clinical information includes medical and demographic data that play a role in diagnosis. As the complexity and volume of medical data increase, the selection of relevant features has become a major challenge in clinical analysis. Conventional methods for feature selection are time-consuming and often result in suboptimal features. Therefore, computer-based automated approaches have emerged as more accurate solutions.*

*The research develops an end-to-end methodology based on metaheuristic algorithms and supervised learning for optimizing feature selection in medical data. The methodology covers the entire process, from raw data processing to model performance evaluation. The research used structured and unstructured data from 11 datasets, employing 10 metaheuristic algorithms and 4 supervised learning algorithms. The results from 157 models show that, for structured data, the models achieved performance above 90% for low-dimensional data and 100% for medium to high-dimensional data across all performance metrics. Meanwhile, for unstructured data, the models achieved 100% accuracy in both binary and multi-class classification.*

*The findings of this research demonstrate that metaheuristic algorithm-based feature selection can enhance the performance of supervised learning models for medical data. Thus, the developed methodology is expected to support more accurate and reliable clinical decision-making.*

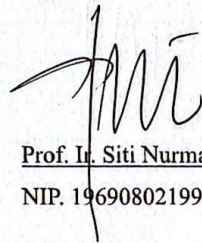
**Keywords:** *metaheuristic algorithms, clinical information, fitness function, supervised learning*

Study Program Coordinator



Prof. Dr. Ir. Nukman, MT  
NIP. 195903211987031001

Supervisor

A blue ink signature of Prof. Ir. Siti Nurmaini, M.T., Ph.D. is written over the text.

Prof. Ir. Siti Nurmaini, M.T., Ph.D  
NIP. 196908021994012001

## DAFTAR ISI

HALAMAN PENGESAHAN.....	iv
HALAMAN PERSETUJUAN .....	iv
PERNYATAAN ORIGINALITAS/PLAGIARISME .....	iv
SURAT KETERANGAN PENGECEKAN SIMILARITY .....	v
ABSTRAK .....	x
ABSTRACT.....	xi
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL .....	xvi
DAFTAR SINGKATAN.....	xvii
KATA PENGANTAR.....	xix
<b>BAB 1 PENDAHULUAN .....</b>	<b>88</b>
1.1. Latar Belakang Masalah.....	88
1.2. Rumusan Masalah .....	89
1.3. Tujuan Penelitian.....	90
1.4. Lingkup Penelitian .....	90
1.5. Sistematika Penulisan.....	91
<b>BAB 2 TINJAUAN PUSTAKA.....</b>	<b>Error! Bookmark not defined.</b>
2.1. Pemilihan fitur.....	<b>Error! Bookmark not defined.</b>
2.2. Algoritma Metaheuristik .....	<b>Error! Bookmark not defined.</b>
2.3. <i>Supervised Learning</i> .....	<b>Error! Bookmark not defined.</b>
<b>BAB 3 METODOLOGI PENELITIAN.....</b>	<b>Error! Bookmark not defined.</b>
3.1. Kerangka Kerja Penelitian.....	<b>Error! Bookmark not defined.</b>
3.2. Persiapan Data.....	<b>Error! Bookmark not defined.</b>
3.3. Pra-pengolahan Data .....	<b>Error! Bookmark not defined.</b>
3.3.1 Data terstruktur.....	<b>Error! Bookmark not defined.</b>
3.3.2 Data tidak terstruktur.....	<b>Error! Bookmark not defined.</b>
3.4. Pemilihan fitur dengan algoritma metaheuristik .....	<b>Error! Bookmark not defined.</b>
3.5. Klasifikasi.....	<b>Error! Bookmark not defined.</b>



3.5.1. <i>Decision Tree</i> .....	<b>Error! Bookmark not defined.</b>
3.5.2. <i>Random Forest</i> .....	<b>Error! Bookmark not defined.</b>
3.5.3. <i>K-Nearest Neighbor</i> .....	<b>Error! Bookmark not defined.</b>
3.5.4. <i>Support Vector Machine</i> .....	<b>Error! Bookmark not defined.</b>
3.6. <i>Fitness Function</i> .....	<b>Error! Bookmark not defined.</b>
<b>BAB 4 HASIL DAN ANALISIS</b> .....	<b>Error! Bookmark not defined.</b>
4.1. Hasil Perancangan Model Algoritma Metaheuristik dan <i>Machine Learning</i> <b>Error! Bookmark not defined.</b>	
4.2. Hasil Optimasi Fitur Pada Data Medis Dengan Variasi Dimensi Fitur .....	<b>Error! Bookmark not defined.</b>
4.2.1. Data Berdimensi Rendah.....	<b>Error! Bookmark not defined.</b>
4.2.2. Data Berdimensi Sedang .....	<b>Error! Bookmark not defined.</b>
4.2.3. Data Berdimensi Tinggi .....	<b>Error! Bookmark not defined.</b>
4.3. Hasil Kinerja Optimasi Fitur dengan <i>Fitness Function</i> ...	<b>Error! Bookmark not defined.</b>
4.3.1. Hasil Kinerja <i>Fitness Function</i> Pada Data Terstruktur	<b>Error! Bookmark not defined.</b>
4.3.2. Hasil Kinerja <i>Fitness Function</i> Pada Data Tidak Terstruktur.....	<b>Error! Bookmark not defined.</b>
4.3.2.1 Klasifikasi NSR dan AF ( <i>Binary classification</i> ) .....	<b>Error! Bookmark not defined.</b>
4.3.2.2 Klasifikasi NSR, AF, SAR, dan SBR ( <i>Multiclass classification</i> ).....	<b>Error! Bookmark not defined.</b>
<b>BAB 5 KESIMPULAN</b> .....	<b>Error! Bookmark not defined.</b>
<b>LUARAN PUBLIKASI</b> .....	<b>Error! Bookmark not defined.</b>
<b>DAFTAR PUSTAKA</b> .....	<b>93</b>
<b>LAMPIRAN</b> .....	<b>Error! Bookmark not defined.</b>
Lampiran 1. <i>Feature significance</i> LD1-DT.....	<b>Error! Bookmark not defined.</b>
Lampiran 2. <i>Feature significance</i> LD1-RF.....	<b>Error! Bookmark not defined.</b>
Lampiran 3. <i>Feature significance</i> LD1-KNN .....	<b>Error! Bookmark not defined.</b>
Lampiran 4. <i>Feature significance</i> LD1-SVM .....	<b>Error! Bookmark not defined.</b>
Lampiran 5. <i>Feature significance</i> LD2-DT.....	<b>Error! Bookmark not defined.</b>
Lampiran 6. <i>Feature significance</i> LD2-RF .....	<b>Error! Bookmark not defined.</b>

Lampiran 7. *Feature significance* LD2-KNN ..... **Error! Bookmark not defined.**  
Lampiran 8. *Feature significance* LD2-SVM ..... **Error! Bookmark not defined.**  
Lampiran 9. *Feature significance* LD3-DT..... **Error! Bookmark not defined.**  
Lampiran 10. *Feature significance* LD3-RF ..... **Error! Bookmark not defined.**  
Lampiran 11. *Feature significance* LD3-KNN ..... **Error! Bookmark not defined.**  
Lampiran 12. *Feature significance* LD3-SVM ..... **Error! Bookmark not defined.**  
Lampiran 13. *Feature significance* MD1-DT ..... **Error! Bookmark not defined.**  
Lampiran 14. *Feature significance* MD1-RF ..... **Error! Bookmark not defined.**  
Lampiran 15. *Feature significance* MD1-KNN ..... **Error! Bookmark not defined.**  
Lampiran 16. *Feature significance* MD1-SVM ..... **Error! Bookmark not defined.**  
Lampiran 17. *Feature significance* MD2-DT ..... **Error! Bookmark not defined.**  
Lampiran 18. *Feature significance* MD2-RF ..... **Error! Bookmark not defined.**  
Lampiran 19. *Feature significance* MD2-KNN ..... **Error! Bookmark not defined.**  
Lampiran 20. *Feature significance* MD2-SVM ..... **Error! Bookmark not defined.**  
Lampiran 21. *Feature significance* MD3-DT ..... **Error! Bookmark not defined.**  
Lampiran 22. *Feature significance* MD3-RF ..... **Error! Bookmark not defined.**  
Lampiran 23. *Feature significance* MD3-KNN ..... **Error! Bookmark not defined.**  
Lampiran 24. *Feature significance* MD3-SVM ..... **Error! Bookmark not defined.**  
Lampiran 25. *Feature significance* HD1-DT ..... **Error! Bookmark not defined.**  
Lampiran 26. *Feature significance* HD1-RF..... **Error! Bookmark not defined.**  
Lampiran 27. *Feature significance* HD1-KNN..... **Error! Bookmark not defined.**  
Lampiran 28. *Feature significance* HD1-SVM..... **Error! Bookmark not defined.**  
Lampiran 29. *Feature significance* HD2-DT ..... **Error! Bookmark not defined.**  
Lampiran 30. *Feature significance* HD2-RF..... **Error! Bookmark not defined.**  
Lampiran 31. *Feature significance* HD2-KNN..... **Error! Bookmark not defined.**  
Lampiran 32. *Feature significance* HD2-SVM..... **Error! Bookmark not defined.**  
Lampiran 33. *Feature significance* HD3-DT ..... **Error! Bookmark not defined.**  
Lampiran 34. *Feature significance* HD3-RF..... **Error! Bookmark not defined.**  
Lampiran 35. *Feature significance* HD3-KNN..... **Error! Bookmark not defined.**  
Lampiran 36. *Feature significance* HD3-SVM..... **Error! Bookmark not defined.**



Lampiran 37. Hasil ekstraksi fitur *time-domain* kelas NSR, AF, SBR dan SAR  
..... **Error! Bookmark not defined.**

## DAFTAR GAMBAR

- Gambar 2.1. Proses Pemilihan Subset Fitur Optimal [14]..... **Error! Bookmark not defined.**
- Gambar 2.2. Diagram alir metode *wrapper* ..... **Error! Bookmark not defined.**
- Gambar 2.3. Kategori metode pemilihan fitur [14] .... **Error! Bookmark not defined.**
- Gambar 2.4. Klasifikasi algoritma metaheuristik ..... **Error! Bookmark not defined.**
- Gambar 2.5. Jumlah algoritma metaheuristik ..... **Error! Bookmark not defined.**
- Gambar 2.6. Sitasi 10 algoritma metaheuristik teratas yang dikutip GS ..... **Error! Bookmark not defined.**
- Gambar 2.7. Diagram alir proses pemilihan fitur *supervised learning* [30]..... **Error! Bookmark not defined.**
- Gambar 3.1. Metodologi penelitian *end-to-end* optimasi pemilihan fitur ..... **Error! Bookmark not defined.**
- Gambar 3.2. Tahapan penelitian optimasi pemilihan fitur ..... **Error! Bookmark not defined.**
- Gambar 3.3. Sampel rekaman NSR, AF, SAR, dan SBR pada sinyal EKG ..... **Error! Bookmark not defined.**
- Gambar 3.4. Sampel plot fitur ekstraksi *frequency-domain* sinyal medis EKG .. **Error! Bookmark not defined.**
- Gambar 3.5. Sampel plot fitur ekstraksi *time-frequency* domain sinyal medis EKG ..... **Error! Bookmark not defined.**
- Gambar 4.1 Hasil perancangan model algoritma metaheuristik TLO ..... **Error! Bookmark not defined.**
- Gambar 4.2. Subset fitur oleh 10 algoritma metaheuristik ..... **Error! Bookmark not defined.**
- Gambar 4.3. Grafik radar dari sepuluh algoritma metaheuristik ..... **Error! Bookmark not defined.**
- Gambar 4.4. Hasil kinerja TLO berdasarkan pengklasifikasi MD1, MD2, dan MD3 ..... **Error! Bookmark not defined.**
- Gambar 4.5. Hasil kinerja TLO berdasarkan pengklasifikasi HD1, HD2, dan HD3 ..... **Error! Bookmark not defined.**
- Gambar 4.6. Hasil 8 pengujian klasifikasi biner NSR dan AF .. **Error! Bookmark not defined.**
- Gambar 4.7. *Heatmap* CM Klasifikasi NSR dan AF .. **Error! Bookmark not defined.**



Gambar 4.8. Hasil pengujian klasifikasi multikelas NSR, AF, SBR dan SAR. ...**Error! Bookmark not defined.**

Gambar 4.9. *Heatmap* CM Klasifikasi NSR, AF, SBR dan SAR.....**Error! Bookmark not defined.**

## DAFTAR TABEL

- Tabel 2.1. Peranan metode *filter*, *wrapper* dan *embedded* pada proses pemilihan fitur ..... **Error! Bookmark not defined.**
- Tabel 2.2. Penelitian-penelitian di lima tahun terakhir ..... **Error! Bookmark not defined.**
- Tabel 2.3. Deskripsi 10 optimasi metaheuristik yang dieksplorasi pada penelitian ..... **Error! Bookmark not defined.**
- Tabel 3.1. Deskripsi *dataset* medis terstruktur dan tidak terstruktur **Error! Bookmark not defined.**
- Tabel 3.2. Jumlah *missing values* pada *dataset* terstruktur ..... **Error! Bookmark not defined.**
- Tabel 3.3. Jumlah total rekaman sinyal EKG NSR, AF, SAR, dan SBR ..... **Error! Bookmark not defined.**
- Tabel 3.4. Statistik *time-domain* berdasarkan RR-interval [77]. **Error! Bookmark not defined.**
- Tabel 3.5. Hasil ekstraksi fitur *frequency-domain* berdasarkan FFT .. **Error! Bookmark not defined.**
- Tabel 3.6. Hasil ekstraksi fitur *time-frequency domain* berdasarkan STFT ..... **Error! Bookmark not defined.**
- Tabel 3.7. Arsitektur AE dengan teknik ekstraksi fitur *deep features* [83] ..... **Error! Bookmark not defined.**
- Tabel 3.8. Arsitektur CNN dengan teknik ekstraksi fitur *deep features* [84] ..... **Error! Bookmark not defined.**
- Tabel 3.9. Parameter – parameter 10 algoritma metaheuristik yang diuji coba... **Error! Bookmark not defined.**
- Tabel 3.10. Parameter – parameter DT yang diuji coba ..... **Error! Bookmark not defined.**
- Tabel 3.11. Parameter – parameter RF yang diuji coba ..... **Error! Bookmark not defined.**
- Tabel 3.12. Parameter – parameter KNN yang diuji coba ..... **Error! Bookmark not defined.**
- Tabel 3.13. Parameter – parameter SVM yang diuji coba ..... **Error! Bookmark not defined.**



- Tabel 4.1. Studi kasus pengujian klasifikasi biner dan multikelas ...**Error! Bookmark not defined.**
- Tabel 4.2. Jumlah fitur yang diekstraksi dan diseleksi klasifikasi biner .....**Error! Bookmark not defined.**
- Tabel 4.3. Subset fitur oleh TLO dalam data dimensi sedang ..**Error! Bookmark not defined.**
- Tabel 4.4. Subset fitur TLO pada data berdimensi tinggi .....**Error! Bookmark not defined.**
- Tabel 4.5. Hasil sampel ekstraksi fitur *time-domain* sinyal .....**Error! Bookmark not defined.**
- Tabel 4.6. Hasil sampel ekstraksi fitur *frequency-domain* sinyal .....**Error! Bookmark not defined.**
- Tabel 4.7. Hasil sampel ekstraksi fitur *time-frequency-domain* sinyal .....**Error! Bookmark not defined.**
- Tabel 4.8. Jumlah fitur yang diekstraksi dan diseleksi klasifikasi multi kelas ....**Error! Bookmark not defined.**
- Tabel 4.9. Hasil sampel ekstraksi fitur *time-domain* sinyal EKG SBR dan SAR **Error! Bookmark not defined.**
- Tabel 4.10. *Benchmark* metode kasus *dataset* medis terstruktur ..... 84
- Tabel 4.11. *Benchmark* metode kasus *dataset* medis tidak terstruktur ..... 87

## DAFTAR SINGKATAN

ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
AE	Autoencoder
AF	Atrial fibrillation
APC	Atrial premature contraction
BBO	Biogeography Based Optimization
CM	Confusion Matrix
CNN	Convolutional neural network
CSA	Cuckoo Search Algorithm
DE	Differential Evolution
DFT	Discrete Fourier Transform
DT	Decision Tree
DWT	Discrete Wavelet Transform
EFHFS	Ensemble-filter-based hybrid Feature Selection
EKG	Elektrokardiogram
FFA	Firefly Algorithm
FFT	Fast Fourier transform
FN	False Negatives
FP	False Positives
GA	Genetic Algorithm
GS	Google Scholar
GWO	Grey Wolf Optimization
HS	Harmony Search
KNN	K-nearest neighbours
LUDB	Lobachevsky University Electrocardiography Database
NB	Naïve Bayes
NSR	Normal Sinus Rhythm

PAC	Premature Atrial Contraction
PVC	Premature Ventricular Contraction
PSO	Particle Swarm Optimization
RBBB	Right Bundle Branch Block
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random forest
SA	Simulated Annealing
SAR	Supraventricular Arrhythmia
SBR	Sinus Bradycardia
STFT	Short-Time fourier transform
SVM	Support vector machine
TLO	Teaching Learning Based Optimization
TN	True Negatives
TP	True Positives
TS	Tabu Search
QTDB	QT Database



## KATA PENGANTAR

Puji syukur penulis panjatkan hadirat Allah SWT, atas limpahan rahmat dan hidayah-Nya dalam penyelesaian Disertasi dengan judul “Optimasi Pemilihan Fitur Berbasis Algoritma Metaheuristik dan *Machine Learning*”. Penelitian ini disusun dalam rangka penyusunan disertasi yang menjadi salah satu persyaratan untuk memperoleh gelar Doktor Ilmu Komputer dari Program Pascasarjana Fakultas Teknik Universitas Sriwijaya. Penulis menyadari bahwa dalam penyusunan penelitian ini masih terdapat kelemahan yang perlu diperkuat dan kekurangan yang perlu dilengkapi.

Dengan tersusunnya penelitian ini, penulis mengucapkan terima kasih dan penghargaan setinggi-tingginya kepada **Prof. Ir. Siti Nurmaini, M.T., Ph.D** selaku Promotor dan **Prof. Dr. Ir. Bambang Tutuko, M.T.** selaku Ko-Promotor yang telah berkenan memberi bimbingan, arahan dan masukan bagi tersusunnya penelitian yang layak untuk disajikan. Penulis juga mengucapkan terima kasih dan penghargaan kepada:

1. Mama, papa, ibu dan ayah selaku orang tua yang telah mendoakan dan memberikan dukungan dalam proses perkuliahan dan penyelesaian disertasi ini.
2. Suami (Iqbal Ramadhan, S. SI) dan kedua anak yang saya cintai (Muhammad Rafanizan Dhanis dan Nazneen Tsabina Dhanis).
3. Kedua saudara kandung (Fitria, S.E dan Muhammad Iqbal S.T) beserta keluarga dalam memberikan dukungan dan doa.
4. Rektor Universitas Sriwijaya, Prof. Dr. Taufiq Marwa, S.E., M.Si
5. Dekan Fakultas Ilmu Komputer Universitas Sriwijaya, Prof. Dr. Erwin, S.Si., M.Si.
6. Dekan Fakultas Teknik Universitas Sriwijaya, Dr. Bhakti Yudho Suprpto, S.T., M.T.
7. Ketua Program Studi Doktor Ilmu Teknik Universitas Sriwijaya, Prof. Dr. Ir. Nukman, M.T.

8. *Intelligent System Research Group* (ISysRG) sebagai grup riset tempat berdiskusi dan menuangkan ide terkait permasalahan disertasi dan infrastruktur lab yang digunakan. Luaran publikasi terkait disertasi ini didanai penuh oleh *Intelligent System Research Group*.
9. Dian Palupi Rini, M.Kom., Ph.D., Dr. Firdaus, S.T., M.Kom., dan Dr. Yesi Novaria Kunang, S.T., M.Kom., selaku dosen penguji internal dan eksternal yang telah memberikan saran dan kritik yang bersifat membangun dalam penyusunan disertasi.
10. *Programmer* riset (Jordan Marcelino dan Rendy Isdwanta) yang sudah berkolaborasi dan berproses pada penelitian ini.
11. Rekan-rekan mahasiswa seangkatan, khususnya Ibu Novi Yusliani dan Ibu Yunita sebagai rekan bertukar pikiran.
12. Bapak/Ibu Dosen Fakultas Ilmu Komputer Universitas Sriwijaya, khususnya Program Studi Sistem Komputer dan Teknik Informatika.
13. Konten video *food vlogger* Ria SW dan drama Korea Selatan (K-drama) yang menjadi ‘tempat rehat’ sejenak dan menjadi salah satu sumber dopamin bagi penulis. Penulis merasa terbantu secara psikologis dalam menjalani proses penelitian disertasi ini.
14. Semua pihak yang telah membantu kegiatan penelitian hingga tersusunnya disertasi ini.

Terima kasih atas perhatiannya.

Palembang, 2025

Annisa Darmawahyuni

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Informasi klinis dapat berupa kumpulan data medis dan data demografi terkait diagnosis dan pengobatan pasien [1]. Ukuran informasi klinis meningkat secara signifikan dalam satu dekade terakhir, khususnya pada data medis akibat munculnya gejala dan tanda klinis baru pada penyakit tertentu [2][3]. Gejala dan tanda klinis dari suatu penyakit menjadi karakteristik khusus, yang jumlahnya dapat mencapai puluhan ribu yang membuat proses analisis menjadi kurang efektif dan tidak akurat [2]. Sebagai contoh, data *microarray* untuk diagnosis tumor dan kanker memiliki karakteristik ribuan gen [3][4], namun tidak semua gen berperan penting dalam pengambilan keputusan klinis [4].

Proses pemilihan karakteristik secara konvensional telah dilakukan secara individual dan tidak menunjukkan kinerja akurat [5]. Hal ini disebabkan sebagian karakteristik khusus data medis perlu dievaluasi secara keseluruhan [5]. Proses pemilihan karakteristik otomatis diperlukan untuk mendukung pengambilan keputusan klinis yang menghasilkan kinerja akurat dalam waktu yang efisien. Semua karakteristik khusus pada data medis dikenal sebagai fitur. Sejumlah fitur akan memberikan informasi penting tentang suatu kondisi klinis [6]. Jika fitur yang dipilih tidak relevan, maka akan menghasilkan informasi klinis tidak akurat [7].

Pemilihan fitur dapat dilakukan melalui dua pendekatan; konvensional dan otomatis berbasis komputer. Pendekatan konvensional membutuhkan waktu yang lama dan fitur suboptimal [8][9][10], sehingga diperlukan pemilihan fitur otomatis berbasis komputer. Proses pemilihan fitur otomatis dibagi dua jenis; (i) melibatkan algoritma *machine learning*, khususnya *supervised learning* yang disebut *dependent classifier*, yaitu *wrapper*, dan (ii) tanpa melibatkan algoritma *supervised learning* disebut *independent classifier*, yaitu *filter*. Metode *wrapper* cenderung berkinerja lebih baik dibanding dengan metode *filter* dari sisi kinerja [6], [11]–[13], selain itu memberikan kombinasi fitur dengan kinerja terbaik pada model prediktif tersebut [3][14].



Algoritma pemilihan fitur dengan pendekatan metode *wrapper* dibagi menjadi tiga jenis; algoritma pemilihan fitur eksponensial [14], berurutan [4], [15], [16] dan acak [14], [17]–[19]. Algoritma pemilihan fitur eksponensial dan berurutan menunjukkan hasil akurat, namun secara teknis sangat sulit diterapkan karena biaya komputasi tinggi [14]. Kedua algoritma tersebut sering mengalami masalah seperti konvergen terlalu dini, sehingga menghasilkan solusi suboptimal, kompleksitas biaya tinggi dan lebih mengarah pada permasalahan lokal optima [14]. Algoritma pemilihan fitur acak lebih unggul dibandingkan algoritma pemilihan fitur eksponensial dan berurutan karena mampu menghindari lokal optima dan meningkatkan kinerja model [14], [20]–[22].

Algoritma pemilihan fitur acak yang optimal adalah algoritma metaheuristik [18][23], yang mampu menemukan subset fitur terbaik dengan menghasilkan solusi paling optimal [18][23]. Algoritma metaheuristik memiliki fleksibilitas tinggi karena mudah diimplementasi [20]. Algoritma metaheuristik dapat dimodifikasi dengan mudah sesuai dengan permasalahan tertentu [18]. Ciri khas utama dari algoritma metaheuristik adalah bahwa metode ini memiliki kemampuan luar biasa menghindari permasalahan konvergen dini [18][23].

Fokus penelitian ini adalah untuk menghasilkan fitur optimal dan meningkatkan kinerja model *supervised learning* pada data medis yang memiliki jumlah fitur bervariasi. Untuk mencapai tujuan tersebut, penelitian ini akan mengembangkan metodologi *end-to-end* untuk pemilihan fitur optimal berbasis algoritma metaheuristik dan *supervised learning* pada data medis. Metodologi *end-to-end* mencakup seluruh proses dari awal (proses masukkan data mentah) hingga akhir (evaluasi hasil kinerja) untuk optimasi pemilihan fitur berbasis algoritma metaheuristik dan *machine learning*. Evaluasi kinerja algoritma metaheuristik dan *supervised learning* dievaluasi untuk menghasilkan *fitness function* terbaik. *Fitness function* yang dicapai adalah fitur optimal dengan kinerja model *supervised learning* akurat dan presisi.

## 1.2. Rumusan Masalah

Beberapa penelitian awal untuk optimasi pemilihan fitur pada data medis yang memiliki dimensi bervariasi dilakukan untuk mengatasi beberapa

permasalahan. Untuk itu, pertanyaan riset yang akan dijawab pada penelitian ini adalah:

*“Bagaimana mengembangkan metodologi end-to-end untuk pemilihan fitur optimal berbasis algoritma metaheuristik dan supervised learning pada data medis?”*

Untuk menyelesaikan permasalahan tersebut, pertanyaan riset di atas terbagi menjadi:

1. Bagaimana merancang model algoritma metaheuristik yang akan diinduksi dengan *supervised learning* pada data medis?
2. Bagaimana mengoptimasi fitur pada data medis yang memiliki dimensi fitur bervariasi?
3. Bagaimana menghasilkan kinerja optimasi fitur dengan *fitness function* terbaik?

### **1.3. Tujuan Penelitian**

Penelitian ini akan mengembangkan metodologi *end-to-end* untuk pemilihan fitur optimal menggunakan algoritma metaheuristik dan *supervised learning*. Untuk mencapai hal tersebut maka tujuan khusus penelitian ini adalah:

1. Merancang model algoritma metaheuristik dan *supervised learning* pada data medis.
2. Mengoptimasi fitur pada data medis yang memiliki dimensi fitur bervariasi.
3. Menghasilkan kinerja optimasi fitur dengan *fitness function* terbaik.

### **1.4. Lingkup Penelitian**

Berikut beberapa batasan penelitian ini adalah:

1. Data yang digunakan pada penelitian ini adalah data medis terstruktur (dimensi rendah, sedang dan tinggi) dan data medis tidak terstruktur.
2. Algoritma metaheuristik yang digunakan dipilih secara acak dari empat kelas, yaitu:
  - a) *Swarm intelligence*: *Artificial Bee Colony* (ABC), *Particle Swarm Optimization* (PSO), *Cuckoo Search Algorithm* (CSA), dan *Firefly Algorithm* (FFA),
  - b) *Evolutionary algorithm*: *Genetic Algorithm* (GA) dan *Differential Evolution* (DE),
  - c) *Physics-based algorithm*: *Harmony Search* (HS), *Simulated Annealing* (SA)
  - d) *Human behaviour-related algorithm*: *Teaching Learning Based Optimization* (TLO/TLBO), *Biogeography Based Optimization* (BBO).
3. Metode *supervised learning* yang digunakan adalah *decision tree* (DT), *random forest* (RF), *k-nearest neighbours* (KNN) dan *support vector machine* (SVM) dikarenakan algoritma tersebut memiliki tingkat bias yang rendah dengan parameter sederhana dan secara umum telah digunakan pada penelitian-penelitian sebelumnya khususnya terkait pemilihan fitur optimal pada data medis.
4. Kinerja yang diimplementasikan untuk memvalidasi model adalah *fitness function*, dengan tujuan untuk menghasilkan fitur optimal dan evaluasi akurasi, sensitivitas, spesifisitas dan presisi.

### **1.5. Sistematika Penulisan**

Proposal penelitian ini terdiri dari lima bab. Sistematika penulisan bertujuan untuk memudahkan penyusunan isi proposal dari setiap bab. Rangkuman proposal penelitian ini terdiri dari:

## **1. BAB I PENDAHULUAN**



Bab ini berisi pendahuluan berupa latar belakang, rumusan masalah, tujuan dan lingkup penelitian yang mengacu pada penelitian-penelitian yang pernah dilakukan sebelumnya.

## **2. BAB II TINJAUAN PUSTAKA**

Bab ini berisi kerangka teori dan studi literatur terkait permasalahan topik yang diusulkan. Bab ini juga menjelaskan tentang penelitian-penelitian sebelumnya yang terkait dengan optimasi pemilihan fitur menggunakan algoritma metaheuristik dan *supervised learning* pada data medis.

## **3. BAB III METODOLOGI PENELITIAN**

Bab ini berisi metodologi yang menjelaskan alur bertahap dan terperinci, tentang langkah-langkah yang digunakan untuk mengumpulkan dan menganalisis permasalahan optimasi pemilihan fitur menggunakan algoritma metaheuristik dan *supervised learning*. Bab ini berisi beberapa tahapan terstruktur untuk mencapai tujuan yang diharapkan.

## **4. BAB IV HASIL DAN ANALISIS**

Bab ini berisi hasil dan analisis dari permasalahan optimasi pemilihan fitur menggunakan algoritma metaheuristik dan *supervised learning*, yang diusulkan berdasarkan tujuan dan lingkup penelitian terkait literatur yang telah diusulkan. Hasil yang diharapkan pada penelitian ini dapat memberikan manfaat pada penelitian selanjutnya.

## **5. BAB V KESIMPULAN**

Bab ini kesimpulan dari metodologi dan pengujian penelitian yang telah dilakukan di tahapan awal penelitian. Bab ini juga menjelaskan keterbatasan penelitian dan pengembangan yang bisa dikembangkan untuk penelitian selanjutnya.

## DAFTAR PUSTAKA

- [1] F. Masood, J. Masood, H. Zahir, K. Driss, N. Mehmood, and H. Farooq, "Novel Approach to Evaluate Classification Algorithms and Feature Selection Filter Algorithms Using Medical Data," *J. Comput. Cogn. Eng.*, vol. 2, no. 1, pp. 57–67, May 2022, doi: 10.47852/bonviewJCCE2202238.
- [2] N. Singh and P. Singh, "A hybrid ensemble-filter wrapper feature selection approach for medical data classification," *Chemom. Intell. Lab. Syst.*, vol. 217, 2021, doi: 10.1016/j.chemolab.2021.104396.
- [3] B. Mirza, W. Wang, J. Wang, H. Choi, N. C. Chung, and P. Ping, "Machine learning and integrative analysis of biomedical big data," *Genes*, vol. 10, no. 2. 2019. doi: 10.3390/genes10020087.
- [4] G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima, and S. M. H. Akhter, "Comparative Study of Optimum Medical Diagnosis of Human Heart Disease Using Machine Learning Technique with and Without Sequential Feature Selection," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3153047.
- [5] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection," *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, 2016, doi: 10.1109/TEVC.2015.2504420.
- [6] B. Remeseiro and V. Bolon-Canedo, "A review of feature selection methods in medical applications," *Computers in Biology and Medicine*, vol. 112. 2019. doi: 10.1016/j.combiomed.2019.103375.
- [7] M. Alirezanejad, R. Enayatifar, H. Motameni, and H. Nematzadeh, "Heuristic filter feature selection methods for medical datasets," *Genomics*, vol. 112, no. 2, 2020, doi: 10.1016/j.ygeno.2019.07.002.
- [8] A. Jamili, "Robust job shop scheduling problem: Mathematical models, exact and heuristic algorithms," *Expert Syst. Appl.*, vol. 55, pp. 341–350, 2016.
- [9] L. E. Sánchez, J. A. Diaz-Pace, and A. Zunino, "A family of heuristic search

- algorithms for feature model optimization,” *Sci. Comput. Program.*, vol. 172, pp. 264–293, 2019.
- [10] J. A. Pereira, L. Maciel, T. F. Noronha, and E. Figueiredo, “Heuristic and exact algorithms for product configuration in software product lines,” *Int. Trans. Oper. Res.*, vol. 24, no. 6, pp. 1285–1306, 2017.
- [11] J. Maldonado, M. C. Riff, and B. Neveu, “A review of recent approaches on wrapper feature selection for intrusion detection,” *Expert Systems with Applications*, vol. 198. 2022. doi: 10.1016/j.eswa.2022.116822.
- [12] W. Liu and J. Wang, “Recursive elimination–election algorithms for wrapper feature selection,” *Appl. Soft Comput.*, vol. 113, 2021, doi: 10.1016/j.asoc.2021.107956.
- [13] S. J. Lee, Z. Xu, T. Li, and Y. Yang, “A novel bagging C4.5 algorithm based on wrapper feature selection for supporting wise clinical decision making,” *J. Biomed. Inform.*, vol. 78, 2018, doi: 10.1016/j.jbi.2017.11.005.
- [14] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, “Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019),” *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3056407.
- [15] H. R. Al Ghayab, Y. Li, S. Abdulla, M. Diykh, and X. Wan, “Classification of epileptic EEG signals based on simple random sampling and sequential feature selection,” *Brain Informatics*, vol. 3, no. 2, 2016, doi: 10.1007/s40708-016-0039-1.
- [16] K. Yan, L. Ma, Y. Dai, W. Shen, Z. Ji, and D. Xie, “Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis,” *Int. J. Refrig.*, vol. 86, 2018, doi: 10.1016/j.ijrefrig.2017.11.003.
- [17] Y. Mao and Y. Yang, “A Wrapper Feature Subset Selection Method Based on Randomized Search and Multilayer Structure,” *Biomed Res. Int.*, vol. 2019, 2019, doi: 10.1155/2019/9864213.
- [18] Z. Wang, X. Xiao, and S. Rajasekaran, “Novel and efficient randomized

- algorithms for feature selection,” *Big Data Min. Anal.*, vol. 3, no. 3, 2020, doi: 10.26599/BDMA.2020.9020005.
- [19] N. Lin, J. Jiang, S. Guo, and M. Xiong, “Functional principal component analysis and randomized sparse clustering algorithm for medical image analysis,” *PLoS One*, vol. 10, no. 7, 2015, doi: 10.1371/journal.pone.0132945.
- [20] L. S. Riza, Iip, E. P. Nugroho, M. B. A. Prabowo, E. Junaeti, and A. G. Abdullah, “Metaheuristicopt: Metaheuristic for optimization,” *R Packag. version 1.0. 0, 2017*, 2019.
- [21] T. M. Le, T. M. Vo, T. N. Pham, and S. V. T. Dao, “A Novel Wrapper-Based Feature Selection for Early Diabetes Prediction Enhanced with a Metaheuristic,” *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2020.3047942.
- [22] S. Kaur, Y. Kumar, A. Koul, and S. Kumar Kamboj, “A Systematic Review on Metaheuristic Optimization Techniques for Feature Selections in Disease Diagnosis: Open Issues and Challenges,” *Archives of Computational Methods in Engineering*, vol. 30, no. 3. 2023. doi: 10.1007/s11831-022-09853-1.
- [23] S. Yildirim, Y. Kaya, and F. Kılıç, “A modified feature selection method based on metaheuristic algorithms for speech emotion recognition,” *Appl. Acoust.*, vol. 173, 2021, doi: 10.1016/j.apacoust.2020.107721.
- [24] J. Zhang, Y. Xiong, and S. Min, “A new hybrid filter/wrapper algorithm for feature selection in classification,” *Anal. Chim. Acta*, vol. 1080, 2019, doi: 10.1016/j.aca.2019.06.054.
- [25] U. M. Khaire and R. Dhanalakshmi, “Stability of feature selection algorithm: A review,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4. 2022. doi: 10.1016/j.jksuci.2019.06.012.
- [26] R. Alanni, J. Hou, H. Azzawi, and Y. Xiang, “Deep gene selection method to select genes from microarray datasets for cancer classification,” *BMC*



- Bioinformatics*, vol. 20, no. 1, 2019, doi: 10.1186/s12859-019-3161-2.
- [27] P. Dhal and C. Azad, “A comprehensive survey on feature selection in the various fields of machine learning,” *Appl. Intell.*, vol. 52, no. 4, 2022, doi: 10.1007/s10489-021-02550-9.
- [28] Z. Xu, I. King, M. R. T. Lyu, and R. Jin, “Discriminative semi-supervised feature selection via manifold regularization,” *IEEE Trans. Neural Networks*, vol. 21, no. 7, 2010, doi: 10.1109/TNN.2010.2047114.
- [29] C. Yan, J. Ma, H. Luo, and A. Patel, “Hybrid binary Coral Reefs Optimization algorithm with Simulated Annealing for Feature Selection in high-dimensional biomedical datasets,” *Chemom. Intell. Lab. Syst.*, vol. 184, 2019, doi: 10.1016/j.chemolab.2018.11.010.
- [30] J. Tang, S. Alelyani, and H. Liu, *Feature selection for classification: A review*. Chapman and Hall/CRC, 2014. doi: 10.1201/b17320.
- [31] A. Jović, K. Brkić, and N. Bogunović, “A review of feature selection methods with applications,” in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, 2015. doi: 10.1109/MIPRO.2015.7160458.
- [32] Z. Sun, G. Bebis, and R. Miller, “Object detection using feature subset selection,” *Pattern Recognit.*, vol. 37, no. 11, 2004, doi: 10.1016/j.patcog.2004.03.013.
- [33] A. K. Jain, R. P. W. Duin, and J. Mao, “Statistical pattern recognition: A review,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, 2000, doi: 10.1109/34.824819.
- [34] K. Hussain, M. N. Mohd Salleh, S. Cheng, and Y. Shi, “Metaheuristic research: a comprehensive survey,” *Artif. Intell. Rev.*, vol. 52, no. 4, 2019, doi: 10.1007/s10462-017-9605-z.
- [35] M. Maghsoudloo, S. Azimzadeh Jamalkandi, A. Najafi, and A. Masoudi-Nejad, “An efficient hybrid feature selection method to identify potential

- biomarkers in common chronic lung inflammatory diseases,” *Genomics*, vol. 112, no. 5, 2020, doi: 10.1016/j.ygeno.2020.06.010.
- [36] R. Abu Khurmaa, I. Aljarah, and A. Sharieh, “An intelligent feature selection approach based on moth flame optimization for medical diagnosis,” *Neural Comput. Appl.*, vol. 33, no. 12, 2021, doi: 10.1007/s00521-020-05483-5.
- [37] M. Mandal, P. K. Singh, M. F. Ijaz, J. Shafi, and R. Sarkar, “A tri-stage wrapper-filter feature selection framework for disease classification,” *Sensors*, vol. 21, no. 16, 2021, doi: 10.3390/s21165571.
- [38] M. Canayaz, “Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods,” *Appl. Soft Comput.*, vol. 128, 2022, doi: 10.1016/j.asoc.2022.109462.
- [39] S. Bashir, I. U. Khattak, A. Khan, F. H. Khan, A. Gani, and M. Shiraz, “A Novel Feature Selection Method for Classification of Medical Data Using Filters, Wrappers, and Embedded Approaches,” *Complexity*, vol. 2022, 2022, doi: 10.1155/2022/8190814.
- [40] M. Mehmood, N. Alshammari, S. A. Alanazi, and F. Ahmad, “Systematic Framework to Predict Early-Stage Liver Carcinoma Using Hybrid of Feature Selection Techniques and Regression Techniques,” *Complexity*, vol. 2022, 2022, doi: 10.1155/2022/7816200.
- [41] M. M. Ershadi and A. Seifi, “Applications of dynamic feature selection and clustering methods to medical diagnosis,” *Appl. Soft Comput.*, vol. 126, 2022, doi: 10.1016/j.asoc.2022.109293.
- [42] M. Alweshah, S. Alkhalaileh, M. A. Al-Betar, and A. A. Bakar, “Coronavirus herd immunity optimizer with greedy crossover for feature selection in medical diagnosis,” *Knowledge-Based Syst.*, vol. 235, 2022, doi: 10.1016/j.knosys.2021.107629.
- [43] A. M. Vommi and T. K. Battula, “A hybrid filter-wrapper feature selection using Fuzzy KNN based on Bonferroni mean for medical datasets classification: A COVID-19 case study,” *Expert Syst. Appl.*, vol. 218, 2023,

doi: 10.1016/j.eswa.2023.119612.

- [44] A. Qtaish, D. Albashish, M. Braik, M. T. Alshammari, A. Alreshidi, and E. J. Alreshidi, "Memory-Based Sand Cat Swarm Optimization for Feature Selection in Medical Diagnosis," *Electron.*, vol. 12, no. 9, 2023, doi: 10.3390/electronics12092042.
- [45] N. Talpur, S. J. Abdulkadir, M. H. Hasan, H. Alhussian, and A. Alwadain, "A Novel Wrapper-Based Optimization Algorithm for the Feature Selection and Classification," *Comput. Mater. Contin.*, vol. 74, no. 3, 2023, doi: 10.32604/cmc.2023.034025.
- [46] R. Shetty, V. S. Bhat, and J. Pujari, "Content-based medical image retrieval using deep learning-based features and hybrid meta-heuristic optimization," *Biomed. Signal Process. Control*, vol. 92, p. 106069, 2024.
- [47] R. R. Mostafa, A. M. Khedr, Z. Al Aghbari, I. Afyouni, I. Kamel, and N. Ahmed, "An adaptive hybrid mutated differential evolution feature selection method for low and high-dimensional medical datasets," *Knowledge-Based Syst.*, vol. 283, p. 111218, 2024.
- [48] S. Agrawal and S. P. Sahu, "Image-based Parkinson disease detection using deep transfer learning and optimization algorithm," *Int. J. Inf. Technol.*, vol. 16, no. 2, pp. 871–879, 2024.
- [49] S. Desale, A. Rasool, S. Andhale, and P. Rane, "Heuristic and Meta-Heuristic Algorithms and Their Relevance to the Real World: A Survey," *Int. J. Comput. Eng. Res. TRENDS*, vol. 351, no. 5, 2015.
- [50] M. Y. Cheng, D. Prayogo, Y. W. Wu, and M. M. Lukito, "A Hybrid Harmony Search algorithm for discrete sizing optimization of truss structure," *Autom. Constr.*, vol. 69, 2016, doi: 10.1016/j.autcon.2016.05.023.
- [51] O. Olorunda and A. P. Engelbrecht, "Measuring exploration/exploitation in particle swarms using swarm diversity," in *2008 IEEE Congress on Evolutionary Computation, CEC 2008*, 2008. doi: 10.1109/CEC.2008.4630938.

- [52] L. Lin and M. Gen, "Auto-tuning strategy for evolutionary algorithms: Balancing between exploration and exploitation," in *Soft Computing*, 2009. doi: 10.1007/s00500-008-0303-2.
- [53] M. Ali, P. Siarry, and M. Pant, "An efficient Differential Evolution based algorithm for solving multi-objective optimization problems," *Eur. J. Oper. Res.*, vol. 217, no. 2, 2012, doi: 10.1016/j.ejor.2011.09.025.
- [54] J. Tang, G. Liu, and Q. Pan, "A Review on Representative Swarm Intelligence Algorithms for Solving Optimization Problems: Applications and Trends," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 10. 2021. doi: 10.1109/JAS.2021.1004129.
- [55] Z. Wei, C. Huang, X. Wang, T. Han, and Y. Li, "Nuclear Reaction Optimization: A Novel and Powerful Physics-Based Algorithm for Global Optimization," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2918406.
- [56] M. Allam and M. Nandhini, "Optimal feature selection using binary teaching learning based optimization algorithm," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 2, 2022, doi: 10.1016/j.jksuci.2018.12.001.
- [57] K. Rajwar, K. Deep, and S. Das, "An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges," *Artif. Intell. Rev.*, 2023, doi: 10.1007/s10462-023-10470-y.
- [58] W. Wu, Y. Huang, and X. Wu, "A New Deep Learning Method with Self-Supervised Learning for Delineation of the Electrocardiogram," *Entropy*, vol. 24, no. 12, p. 1828, 2022.
- [59] A. Madani, J. R. Ong, A. Tibrewal, and M. R. K. Mofrad, "Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease," *NPJ Digit. Med.*, vol. 1, no. 1, pp. 1–11, 2018.
- [60] S. Zhou, Q. Chen, and X. Wang, "Active deep learning method for semi-supervised sentiment classification," *Neurocomputing*, vol. 120, pp. 536–



546, 2013.

- [61] S. Hadiyoso, S. Aulia, I. D. Irawati, and M. Ramdhani, “Multi-Distance Dispersion Entropy for ECG Signal Classification,” *Int. J. online Biomed. Eng.*, vol. 18, no. 7, 2022, doi: 10.3991/ijoe.v18i07.30055.
- [62] M. Chae, S. Han, H. Gil, N. Cho, and H. Lee, “Prediction of in-hospital cardiac arrest using shallow and deep learning,” *Diagnostics*, vol. 11, no. 7, 2021, doi: 10.3390/diagnostics11071255.
- [63] C. C. Aggarwal, *Data classification: algorithms and applications*. CRC press, 2014.
- [64] J. Ha, M. Kambe, and J. Pe, *Data Mining: Concepts and Techniques*. 2011. doi: 10.1016/C2009-0-61819-5.
- [65] S. Kimura and H. Yabe, “UCI Machine Learning Repository. Irvine, CA : University of California, School of Information and Computer Science,” <http://archive.ics.uci.edu/ml>, vol. 55, no. 0, 2010.
- [66] D. Sculley, J. Elliott, and J. Moser, “Kaggle Dataset.” Accessed: Mar. 02, 2023. [Online]. Available: <https://www.kaggle.com/datasets>
- [67] R. C. Pereira, M. S. Santos, P. P. Rodrigues, and P. H. Abreu, “Reviewing autoencoders for missing data imputation: Technical trends, applications and outcomes,” *J. Artif. Intell. Res.*, vol. 69, pp. 1255–1285, 2020.
- [68] A. Darmawahyuni *et al.*, “Improved delineation model of a standard 12-lead electrocardiogram based on a deep learning algorithm,” *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, p. 139, 2023.
- [69] S. Nurmaini *et al.*, “An Automated ECG Beat Classification System Using Deep Neural Networks with an Unsupervised Feature Extraction Technique,” *Appl. Sci.*, vol. 9, no. 14, pp. 1–17, 2019, doi: 10.3390/app9142921.
- [70] S. Nurmaini *et al.*, “Robust electrocardiogram delineation model for automatic morphological abnormality interpretation,” *Sci. Rep.*, vol. 13, no.

- 1, p. 13736, 2023.
- [71] A. K. Singh and S. Krishnan, “ECG signal feature extraction trends in methods and applications,” *Biomed. Eng. Online*, vol. 22, no. 1, p. 22, 2023.
- [72] S. Chandra, A. Sharma, and G. K. Singh, “Feature extraction of ECG signal,” *J. Med. Eng. Technol.*, vol. 42, no. 4, pp. 306–316, 2018.
- [73] S. Kutscher, “Algorithms for ECG Feature Extraction: an Overview,” 2013.
- [74] R. Nanjundegowda and V. A. Meshram, “Arrhythmia Detection Based on Hybrid Features of T-wave in Electrocardiogram,” *Int. J. Intell. Eng. Syst.*, vol. 11, no. 1, pp. 153–162, 2018.
- [75] M. Kropf, D. Hayn, and G. Schreier, “ECG classification based on time and frequency domain features using random forests,” in *2017 Computing in Cardiology (CinC)*, 2017, pp. 1–4.
- [76] K. Fox *et al.*, “Resting Heart Rate in Cardiovascular Disease,” *Journal of the American College of Cardiology*, vol. 50, no. 9. 2007. doi: 10.1016/j.jacc.2007.04.079.
- [77] A. Parsi, D. Byrne, M. Glavin, and E. Jones, “Heart rate variability feature selection method for automated prediction of sudden cardiac death,” *Biomed. Signal Process. Control*, vol. 65, p. 102310, 2021.
- [78] P. M. Tripathi, A. Kumar, R. Komaragiri, and M. Kumar, “Watermarking of ECG signals compressed using Fourier decomposition method,” *Multimed. Tools Appl.*, vol. 81, no. 14, pp. 19543–19557, 2022.
- [79] R. M. Rangayyan and S. Krishnan, *Biomedical signal analysis*. John Wiley & Sons, 2024.
- [80] R. Abeysekera and B. Boashash, “Time-frequency domain features of ECG signals: their application in P wave detection using the cross Wigner-Ville distribution,” in *International Conference on Acoustics, Speech, and Signal Processing*, 1989, pp. 1524–1527.
- [81] C. Mateo and J. A. Talavera, “Short-time Fourier transform with the window

- size fixed in the frequency domain,” *Digit. Signal Process.*, vol. 77, pp. 13–21, 2018.
- [82] V. X. Afonso and W. J. Tompkins, “Detecting ventricular fibrillation,” *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 152–159, 1995.
- [83] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, “Deep-ECG: Convolutional neural networks for ECG biometric recognition,” *Pattern Recognit. Lett.*, 2018.
- [84] V. V Kuznetsov, V. A. Moskalenko, D. V Gribanov, and N. Y. Zolotykh, “Interpretable feature generation in ECG using a variational autoencoder,” *Front. Genet.*, vol. 12, p. 638191, 2021.
- [85] A. Mumuni and F. Mumuni, “Automated data processing and feature engineering for deep learning and big data applications: a survey,” *J. Inf. Intell.*, 2024.
- [86] H. T. Ibrahim, W. J. Mazher, O. N. Uçan, and O. Bayat, “Feature selection using salp swarm algorithm for real biomedical datasets,” 2017.
- [87] M. N. Anyanwu and S. G. Shiva, “Comparative Analysis of Serial Decision Tree Classification Algorithms,” *Int. J. Comput. Sci. Secur.*, vol. 3, no. 3, 2009.
- [88] T. Daniya, M. Geetha, and K. S. Kumar, “Classification and regression trees with gini index,” *Adv. Math. Sci. J.*, vol. 9, no. 10, 2020, doi: 10.37418/amsj.9.10.53.
- [89] M. Belgiu and L. Drăgu, “Random forest in remote sensing: A review of applications and future directions,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, 2016. doi: 10.1016/j.isprsjprs.2016.01.011.
- [90] Y. Liu, Y. Wang, and J. Zhang, “New machine learning algorithm: Random forest,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012. doi: 10.1007/978-3-642-34062-8\_32.

- [91] K. Taunk, S. De, S. Verma, and A. Swetapadma, “A brief review of nearest neighbor algorithm for learning and classification,” in *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, 2019. doi: 10.1109/ICCS45141.2019.9065747.
- [92] C.-W. Hsu and C.-J. Lin, “A comparison of methods for multiclass support vector machines,” *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.
- [93] R. Sharma, P. Madan, S. Hariharan, and S. Mahajan, “Hybrid Radial Basis Function and Support Vector Machine Model for Precise Breast Cancer Diagnosis,” in *2024 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 2024, pp. 35–38.
- [94] M. Soleimani, A. Harooni, N. Erfani, A. R. Khan, T. Saba, and S. A. Bahaj, “Classification of cancer types based on microRNA expression using a hybrid radial basis function and particle swarm optimization algorithm,” *Microsc. Res. Tech.*, vol. 87, no. 5, pp. 1052–1062, 2024.
- [95] R. Jain, V. Kukreja, S. Chattopadhyay, A. Verma, and R. Sharma, “Radial Basis Function Integrated with Support Vector Machine Model for Breast Cancer Detection,” in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, 2024, pp. 1–5.
- [96] N. Van Thieu and S. Mirjalili, “MEALPY: An open-source library for latest meta-heuristic algorithms in Python,” *J. Syst. Archit.*, vol. 139, 2023, doi: 10.1016/j.sysarc.2023.102871.
- [97] A. C. Cinar, “A comprehensive comparison of accuracy-based fitness functions of metaheuristics for feature selection,” *Soft Comput.*, vol. 27, no. 13, pp. 8931–8958, 2023, doi: 10.1007/s00500-023-08414-3.
- [98] N. Al-Madi, H. Faris, and S. Mirjalili, “Binary multi-verse optimization algorithm for global optimization and discrete problems,” *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 12, 2019, doi: 10.1007/s13042-019-00931-8.

- [99] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, 2016, doi: 10.1016/j.neucom.2016.03.101.
- [100] S. Nurmaini *et al.*, "Deep Learning-Based Stacked Denoising and Autoencoder for ECG Heartbeat Classification," *Electronics*, vol. 9, no. 1, p. 135, 2020, doi: 10.3390/electronics9010135.
- [101] J. Cai, G. Zhou, M. Dong, X. Hu, G. Liu, and W. Ni, "Real-time arrhythmia classification algorithm using time-domain ECG feature based on FFNN and CNN," *Math. Probl. Eng.*, vol. 2021, pp. 1–17, 2021.
- [102] Y. Zhang, J. Yi, A. Chen, and L. Cheng, "Cardiac arrhythmia classification by time--frequency features inputted to the designed convolutional neural networks," *Biomed. Signal Process. Control*, vol. 79, p. 104224, 2023.
- [103] B. Wang *et al.*, "Arrhythmia disease diagnosis based on ECG time--frequency domain fusion and convolutional neural network," *IEEE J. Transl. Eng. Heal. Med.*, vol. 11, pp. 116–125, 2022.
- [104] A. Kumar M and A. Chakrapani, "Classification of ECG signal using FFT based improved Alexnet classifier," *PLoS One*, vol. 17, no. 9, p. e0274225, 2022.
- [105] A. Darmawahyuni *et al.*, "An improved electrocardiogram arrhythmia classification performance with feature optimization," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, p. 412, 2024.