



**OPTIMASI RANDOM FOREST UNTUK DIAGNOSIS PENYAKIT  
GINJAL KRONIK DENGAN MENGGUNAKAN PARTICLE SWARM  
OPTIMIZATION**

**SKRIPSI**

**SHEVA NAUFALRIFQI**

**1810511080**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**2022**



**OPTIMASI RANDOM FOREST UNTUK DIAGNOSIS PENYAKIT  
GINJAL KRONIK DENGAN MENGGUNAKAN PARTICLE SWARM  
OPTIMIZATION**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana  
Komputer**

**SHEVA NAUFALRIFQI**

**1810511080**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**2022**

## PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan sumber yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Sheva NaufalRifqi

NIM : 181011080

Tanggal : 23 Juni 2022

Sebagai persyaratan yang diperlukan untuk mengikuti ujian Sidang Tugas Akhir/Skripsi pada Program Studi S1 Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Depok, 23 Juni 2022



Yang menyatakan,

Sheva NaufalRifqi

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK  
KEPENTINGAN AKADEMIS**

---

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta,  
Saya yang bertanda tangan dibawah ini :

Nama : Sheva NaufalRifqi  
NIM : 1810511080  
Fakultas : Ilmu Komputer  
Program Studi : S1 Informatika

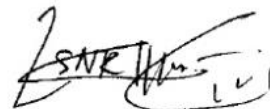
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah Saya yang berjudul:

Optimasi *Random Forest* Untuk Diagnosis Penyakit Ginjal Kronik Dengan Menggunakan *Particle Swarm Optimization*. Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi Saya selama tetap mencantumkan nama Saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Dibuat di : Depok

Pada Tanggal : 23 Juni 2022

Yang Menyatakan,



Sheva NaufalRifqi

## PENGESAHAN

Skripsi diajukan oleh :

Nama : Sheva NaufalRifqi  
NIM : 1810511080  
Program Studi : S1 Informatika  
Judul Skripsi : Optimasi Random Forest Untuk Diagnosis Penyakit Ginjal Kronik  
Menggunakan Particle Swarm Optimization

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widiastiwi, S.Kom, M.Si

Penguji 1



Mayanda Mega Santoni, S.Kom, M.Kom

Penguji 2



Anita Muliawati, S.Kom, M.TI

Pembimbing 1



Desta Sandya Prasvita, S.Kom, M.Kom

Pembimbing 2



Dr. Ermatita, M. Kom.

Dekan



Desta Sandya Prasvita, S.Kom, M.Kom

Kaprodi

Ditetapkan di : Jakarta  
Tanggal Ujian : Selasa, 12 Juli 2022



## HALAMAN PERSETUJUAN

Dengan ini menyatakan bahwa skripsi/tugas akhir berikut:

Nama : Sheva NaufalRifqi  
NIM : 181011080  
Program Studi : Informatika  
Judul : Optimasi Random Forest untuk Diagnosa Penyakit Ginjal Kronis dengan Menggunakan Particle Swarm Optimization

Sebagai persyaratan yang diperlukan untuk mengikuti ujian Sidang Skripsi pada Program Studi S1 Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Mengetahui,



Kaprodi Informatika

Desta Sandya P., S.Kom, M.Kom

NIP/NIDN : 0303128702

Menyetujui,



Dosen Pembimbing 1

Anita Muliawati, S.Kom, M.TI

NIP/NIDN : 0321057001



Dosen Pembimbing 2

Desta Sandya P., S.Kom, M.Kom

NIP/NIDN : 0303128702

Ditetapkan : Jakarta

Tanggal Persetujuan : 23 Juni 2022

# OPTIMASI RANDOM FOREST UNTUK DIAGNOSIS PENYAKIT GINJAL KRONIK DENGAN MENGGUNAKAN PARTICLE SWARM OPTIMIZATION

Sheva NaufalRifqi

## ABSTRAK

Penyakit Ginjal Kronik (PGK) merupakan salah satu penyakit yang penderitanya terus meningkat dalam skala global. Penyakit ginjal kronik ini menyebabkan kemampuan cairan elektrolit pada tubuh tidak dapat mempertahankan metabolisme dalam tubuh dengan baik. Penyebab penyakit ini terus meningkat dikarenakan sifatnya yang sangat *progresif* dan *irreversible*. Untuk mengatasi hal ini, diperlukannya metode cepat dan akurat dalam mendiagnosa penyakit ginjal kronik ini, agar penanganan terhadap penderitanya bisa cepat ditangani. Salah satu metode yang tepat dalam memprediksi diagnosa ginjal kronik ini adalah dengan membangun model klasifikasi dengan menggunakan berbagai macam algoritma, salah satunya dengan menggunakan *random forest*. Algoritma *Random Forest* ini banyak digunakan dalam membangun model klasifikasi, akan tetapi dalam penerapannya diperlukan metode lain untuk mengoptimasi algoritma tersebut agar menjadi lebih akurat. Untuk mengatasi hal tersebut digunakan algoritma *Particle Swarm Optimization* untuk dilakukan seleksi fitur terhadap data yang memiliki fitur yang banyak. Hasil evaluasi dalam pengujian performa dengan menggunakan *Particle Swarm Optimization* dalam pengklasifikasian CKD dan Non CKD menghasilkan akurasi sebesar 99.167%. Seleksi fitur tersebut terbukti sangat efektif dalam optimasinya, karena fitur yang telah diseleksi membuat data lebih optimal untuk diolah dan menghasilkan akurasi yang lebih baik.

**Keyword:** ginjal kronik, *random forest*, *particle swarm optimization*

# OPTIMASI RANDOM FOREST UNTUK DIAGNOSIS PENYAKIT GINJAL KRONIK DENGAN MENGGUNAKAN PARTICLE SWARM OPTIMIZATION

**Sheva NaufalRifqi**

## ABSTRACT

Chronic Kidney Disease (CKD) is one of the diseases whose sufferers continue to increase on a global scale. This chronic kidney disease causes the ability of the electrolyte fluid in the body to not be able to maintain metabolism in the body properly. The causes of this disease continue to increase because it is highly progressive and irreversible. To overcome this, we need a method that is fast and accurate in diagnosing kidney disease, so that the treatment of the sufferer can be handled quickly. One of the appropriate methods in predicting the diagnosis of chronic kidney disease is to build a classification model using various algorithms, one of which is by using a random forest. This Random Forest algorithm is widely used in building classification models, but in its application other methods are needed to optimize the algorithm to be more accurate. To overcome this problem, the Particle Swarm Optimization algorithm is used to perform feature selection on data that has many features. The results of the evaluation in performance testing using Particle Swarm Optimization in classifying CKD and Non CKD, the quality of accuracy is 99.167%. The choice of features has proven to be very effective in its optimization, because the selected features make the data more optimal for processing and produce better accuracy.

**Keyword:** chronic kidney, *random forest*, *particle swarm optimization*



## **KATA PENGANTAR**

Puji serta syukur penulis panjatkan kehadirat Allah SWT atas segala limpahan rahmat, karunia dan ridho-Nya, sehingga Skripsi yang berjudul “Optimasi Random Forest untuk Diagnosa Penyakit Ginjal Kronis dengan Menggunakan Particle Swarm Optimization” berhasil diselesaikan. Dan ucapan terima kasih yang ingin penulis ucapkan kepada :

1. Kedua orang tua penulis, Syaiful (Bapak) dan Nina Hestin Dahrina (Ibu), serta anggota keluarga lain yang selalu memberikan doa dan motivasi sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Desta Sandya P., S.Kom, M.Kom dan Ibu Anita Muliawati, S.Kom, M.TI, selaku dosen pembimbing yang telah membimbing serta memberikan saran serta feedback yang membangun.
3. Ibu Dr. Ermatita, M. Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
4. Bapak Desta Sandya P., S.Kom, M.Kom. selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
5. Bapak Desta Sandya P., S.Kom, M.Kom selaku dosen pembimbing akademik.
6. Bapak/Ibu dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta yang telah memberikan ilmunya kepada Saya.
7. Teman-teman Informatika angkatan 2018, dan rekan sepergaulan selama Saya berkuliah di UPN Veteran Jakarta. Terima kasih atas pengalaman, dan bantuannya ketika mengalami kesulitan, terutama Arvi Arkadia.
8. Terima kasih juga atas segala dukungan partner Saya, Adzani Indah Utami yang telah memberikan bantuan serta doanya, serta teman – teman sekalian yang tidak bisa saya sebutkan satu – persatu.

Akhir kata penulis ucapkan terima kasih dan semoga skripsi ini dapat bermanfaat bagi pembaca.

Depok, 23 Juni 2022

Penulis

## DAFTAR ISI

<b>PERNYATAAN ORISINALITAS</b> .....	ii
<b>PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS</b> .....	iii
<b>PENGESAHAN</b> .....	iv
<b>HALAMAN PERSETUJUAN</b> .....	v
<b>ABSTRAK</b> .....	ii
<b>ABSTRACT</b> .....	iii
<b>KATA PENGANTAR</b> .....	iv
<b>DAFTAR ISI</b> .....	v
<b>DAFTAR GAMBAR</b> .....	viii
<b>DAFTAR TABEL</b> .....	ix
<b>DAFTAR LAMPIRAN</b> .....	x
<b>BAB I PENDAHULUAN</b> .....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	2
1.3. Ruang Lingkup .....	3
1.4. Tujuan Penelitian.....	3
1.5. Luaran Yang Diharapkan .....	3
1.7. Sistematika Penulisan.....	4
<b>BAB II LANDASAN TEORI</b> .....	5
2.1 Penyakit Ginjal Kronis .....	5
2.2 Praproses Data .....	5
2.3 Algoritma Random Forest .....	7
2.3.1. Implementasi Random Forest .....	7

2.3.2.	Bootstrap .....	8
2.4	Algoritma Particle Swarm Optimization .....	9
2.4.1.	Implementasi Particle Swarm Optimization .....	10
2.4.2.	Seleksi Fitur .....	12
2.5	Confusion Matrix .....	13
2.6	Penelitian Terdahulu.....	15
<b>BAB III METODE PENELITIAN .....</b>		<b>16</b>
3.1	Kerangka Pikir.....	16
3.2	Studi Pustaka .....	16
3.3	Pemahaman Data .....	17
3.4	Praproses Data.....	18
3.5	Pencarian Model Terbaik Menggunakan GridSearchCV.....	18
3.6	Penerapan Algoritma Optimasi .....	18
3.7	Seleksi Fitur.....	19
3.8	Pengujian .....	20
3.9	Evaluasi .....	20
3.8	Alat Bantu Penelitian.....	20
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>22</b>
4.1	Pengumpulan Data .....	22
4.2	Hasil dan Implementasi Program Praproses Data .....	22
4.2.1.	Pengolahan Data yang Hilang.....	22
4.2.2.	Mapping Terhadap Data String.....	22
4.2.3.	Mengisi Dataset yang Hilang dengan Nilai Median .....	23
4.3	Membangun Model Random Forest.....	25
4.4	Implementasi Random Forest.....	26

4.5	Implementasi Particle Swarm Optimization pada Model Random Forest	27
4.5.1.	Implementasi PSO dengan Random Forest .....	28
4.5.2.	Fitur yang Terseleksi.....	29
4.5.3.	Penerapan Algoritma <i>Random Forest</i> Terhadap Fitur yang Terseleksi.....	31
4.5	Hasil Perbandingan Akurasi Optimasi PSO dengan Random Forest.....	32
4.6	Perbandingan Metode dengan Penelitian Terdahulu.....	33
<b>BAB V PENUTUP</b> .....		35
5.1	Kesimpulan.....	35
5.2	Saran.....	35
<b>DAFTAR PUSTAKA</b> .....		36
<b>RIWAYAT HIDUP</b> .....		39

## DAFTAR GAMBAR

Gambar 2.1. Ilustrasi Metode Bootstrap .....	9
Gambar 3.1. Alur Kerangka Pikir .....	16
Gambar 4.1. Source Code untuk Implementasi Penggantian Data Hilang dengan Nilai NaN .....	22
Gambar 3.2. Source Code untuk Implementasi Mapping Terhadap Nilai <i>Boolean</i> Menjadi Nilai Biner.....	23
Gambar 4.3. Source Code untuk Implementasi Penggantian Data Hilang dengan Nilai Median.....	23
Gambar 4.4. Source Code untuk Melakukan Drop Terhadap Atribut yang Tidak Terpakai.....	24
Gambar 4.5. Source Code Implementasi <i>Random Forest</i> yang Telah Menggunakan Model Optimal .....	26
Gambar 4.6. Hasil <i>Classification Report</i> .....	26
Gambar 4.7. <i>Plot Confusion Matrix</i> .....	27
Gambar 4.8. Source Code Pendefinisian Method Problem Pada <i>Random Forest</i> dan <i>Method</i> Seleksi Fitur yang Dilakukan .....	28
Gambar 4.9. Source Code Implementasi <i>Particle Swarm Optimization</i> .....	29
Gambar 4.10. Hasil Perbandingan <i>Scores</i> Tiap Fitur.....	30
Gambar 4.11. <i>Source Code</i> Penerapan <i>Random Forest Classifier</i> Terhadap Fitur yang Telah Terseleksi .....	31
Gambar 4.12. Hasil <i>Classification Report Random Forest</i> .....	32
Gambar 4.13. Hasil <i>Classification Report Random Forest</i> Tanpa PSO .....	32

## DAFTAR TABEL

Tabel 2.1. <i>Confusion Matrix</i> .....	14
Tabel 3.1. Atribut Data.....	17
Tabel 4.1. Atribut Data Setelah Pengolahan .....	23
Tabel 4.2. Pembagian Data <i>Training</i> dan Data <i>Testing</i> .....	25
Tabel 4.3. Hasil Pengujian Parameter .....	25
Tabel 4.4. Hasil Model <i>Random Forest</i> Terbaik.....	26
Tabel 4.5. Hasil <i>Scores</i> Tiap Fitur .....	29
Tabel 4.6. Fitur yang Terseleksi.....	31
Tabel 4.7. Perbandingan Akurasi Fitur yang Terseleksi dengan Semua Fitur.....	32
Tabel 4.8. Perbandingan Akurasi dengan PSO dan Tanpa PSO .....	32
Tabel 4.9. Perbandingan Akurasi dengan Penelitian Terdahulu .....	33

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Hasil Turnit in .....	41
--	----