

**OPTIMASI RANDOM SEARCH MENGGUNAKAN  
ALGORITMA RANDOM FOREST UNTUK PREDIKSI  
PENYAKIT LIVER**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh  
**RIYAN BAYU SATRIYA**  
**21.11.3885**

Kepada  
**FAKULTAS ILMU KOMPUTER**  
**UNIVERSITAS AMIKOM YOGYAKARTA**  
**YOGYAKARTA**  
**2025**

**OPTIMASI RANDOM SEARCH MENGGUNAKAN  
ALGORITMA RANDOM FOREST UNTUK PREDIKSI  
PENYAKIT LIVER**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh  
**RIYAN BAYU SATRIYA**  
**21.11.3885**

Kepada  
**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2025**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**OPTIMASI RANDOM SEARCH MENGGUNAKAN ALGORITMA  
RANDOM FOREST UNTUK PREDIKSI PENYAKIT LIVER**

yang disusun dan diajukan oleh

**RIYAN BAYU SATRIYA**

**21.11.3885**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 21 April 2025

**Dosen Pembimbing,**



**Kusnawi, S.Kom, M.Eng**  
**NIK. 190302112**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**OPTIMASI RANDOM SEARCH MENGGUNAKAN ALGORITMA**  
**RANDOM FOREST UNTUK PREDIKSI PENYAKIT LIVER**

yang disusun dan diajukan oleh

**RIYAN BAYU SATRIYA**

**21.11.3885**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 21 April 2025

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Bayu Setiaji, M.Kom**  
**NIK. 190302216**

**Rumini, S.Kom., M.Kom.**  
**NIK. 190302246**

**Kusnawi, S.Kom., M.Eng.**  
**NIK. 190302112**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 21 April 2025

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusrini, S.Kom., M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : RIYAN BAYU SATRIYA  
NIM : 21.11.3885

Menyatakan bahwa Skripsi dengan judul berikut:

**Optimasi Random Search Menggunakan Algoritma Random Forest Untuk  
Prediksi Penyakit Liver**

Dosen Pembimbing : Kusnawi, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 April 2025

Yang Menyatakan,



RIYAN BAYU SATRIYA

## MOTTO

*"Lamun Siro Sekti, Ojo Mateni Lamun Siro Banter, Ojo Ndhisiki Lamun Siro  
Pinter, Ojo Minteri"*

(meskipun kamu sakti (berkuasa) jangan sekali-kali menjatuhkan, meskipun kamu  
cepat jangan selalu mendahului, Meskipun kamu pintar, jangan sok pintar.)

*"Siput memang lambat tapi ia tak pernah berjalan mundur"*

*"manusia bebas untuk mencintai tapi tak bebas untuk memiliki"*



## HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kepada Allah SWT atas segala limpahan rahmat dan ridhoNya yang telah memberikan kesehatan, kelancaran dan kekuatan. Atas segala karunia serta kemudahan yang Engkau berikan akhirnya skripsi ini dapat terselesaikan. Selain itu penulis juga berterimakasih kepada orang yang sangat berarti dalam pembuatan skripsi ini:

1. Kedua Orang Tua, Bapak Sumarli dan Ibu Jumini Yuni Astuti, adik tercinta, Titiya Dyah Gayatri, yang telah memberikan doa, motivasi, semangat yang tak ternilai harganya.
2. Dosen pembimbing, Bapak Kusnawi, S.Kom, M.Eng., yang telah membimbing dari awal hingga terselesaikannya skripsi.
3. Teman-teman SMK N 1 BOYOLALI (TKJ1 Angkatan 2020), Teman-teman Universitas Amikom Yogyakarta (IF01 Angkatan 2021), Teman-teman Universitas Teknologi Yogyakarta dan Kakak-kakak Purbo art yang sudah banyak membantu saya dalam proses pengerjaan skripsi.



## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT atas karunia dan limpahan rahmat dan ridho-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul "OPTIMASI RANDOM SEARCH MENGGUNAKAN ALGORITMA RANDOM FOREST UNTUK PREDIKSI PENYAKIT LIVER". Serta tidak lupa juga shalawat beserta salam kepada baginda Nabi Muhammad shallallahu 'alaihi wasallam beserta keluarganya, para sahabatnya dan ahli sunnah.

Skripsi ini ditulis untuk memenuhi persyaratan untuk mencapai gelar Sarjana Komputer dalam Program Studi Informatika Fakultas Ilmu Komputer. Selama penulisan skripsi ini, penulis telah banyak mendapatkan bimbingan dan bantuan dari berbagai pihak, baik berupa kritik, saran, motivasi serta bantuan lainnya. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua Orang Tua, Bapak Sumarli dan Ibu Jumini Yuni Astuti, adik tercinta, Titiya Dyah Gayatri, yang telah memberikan doa, motivasi, semangat yang tak ternilai harganya.
2. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas AMIKOM Yogyakarta.
3. Bapak Prof. Dr. Kusrini, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
4. Ibu Eli Pujiastuti, M.Kom, selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
5. Bapak Kusnawi, S.Kom, M.Eng, selaku dosen pembimbing penulis yang telah memberikan kritik, saran, waktu, motivasi dan bimbingan dalam skripsi ini sehingga skripsi ini dapat selesai dengan baik dan tepat waktu.



6. Bapak Bayu Setiaji, M.Kom., dan ibu Rumini, S.Kom., selaku dosen penguji yang telah memberikan saran dan masukan atas penyelesaian skripsi ini.
7. Bapak dan Ibu Dosen Program Studi Informatika Universitas AMIKOM Yogyakarta yang telah mengajari penulis dari semester awal hingga akhir sehingga penulis memperoleh banyak sekali ilmu yang bermanfaat dan pengalaman yang sangat berharga.
8. Teman-teman SMK N 1 BOYOLALI (TKJ1 Angkatan 2020), Teman-teman Universitas Amikom Yogyakarta (IF01 angkatan 2021), Teman-teman Universitas Teknologi Yogyakarta dan Kakak-kakak Purbo\_art yang sudah banyak membantu saya dalam proses pengerjaan skripsi.

Yogyakarta, 21 April 2025



RIYAN BAYU SATRIYA

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN.....	xiv
DAFTAR LAMBANG DAN SINGKATAN .....	xv
DAFTAR ISTILAH.....	xvi
INTISARI .....	xviii
ABSTRACT.....	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA .....	5
2.1 Studi Literatur .....	5

2.2	Dasar Teori .....	12
2.2.1	Penyakit Liver .....	12
2.2.2	Machine Learning .....	14
2.2.3	EDA .....	15
2.2.4	Preprocessing Data .....	15
2.2.5	SMOTE .....	16
2.2.6	Random Forest .....	16
2.2.7	Random Search .....	18
2.2.8	Evaluasi .....	19
<b>BAB III METODE PENELITIAN .....</b>		<b>22</b>
3.1	Objek Penelitian .....	22
3.2	Alur Penelitian .....	22
3.2.1	Pengumpulan Data .....	23
3.2.2	EDA .....	24
3.2.3	Preprocessing Data .....	25
3.2.4	SMOTE .....	25
3.2.5	Split Data .....	26
3.2.6	Random Forest .....	26
3.2.7	Random Search .....	29
3.2.8	Evaluasi .....	30
3.3	Alat dan Bahan .....	30
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>32</b>
4.1	Hasil Pembahasan .....	32
4.2	Pengumpulan Data .....	32
4.3	EDA .....	33

4.4 Preprocessing Data.....	35
4.4.1 Pengecekan Heatmap atau Correlation Matrix .....	35
4.4.2 Feature Selection.....	37
4.5 SMOTE .....	41
4.6 Split Data .....	46
4.7 Model Random Forest (Tanpa Optimasi) .....	48
4.8 Optimasi Model Random Forest Menggunakan Random Search.....	52
4.9 Evaluasi .....	59
4.10 Hasil Deployment .....	61
BAB V PENUTUP .....	63
5.1 Kesimpulan.....	63
5.2 Saran.....	63
REFERENSI .....	65
LAMPIRAN.....	71

## DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian .....	7
Tabel 2. 2 Confusion Matrix .....	19
Tabel 3. 1 Deskripsi Fitur Dataset Liver .....	23
Tabel 3. 2 Pembagian Data .....	26
Tabel 3. 3 Parameter Random Search .....	29
Tabel 3. 4 Spesifikasi Hardware Software .....	30
Tabel 4. 1 Pengumpulan Data .....	32
Tabel 4. 2 EDA Informasi struktur dataset diperoleh menggunakan dataset.info()	33
Tabel 4. 3 Menampilkan Statistik Deskriptif Dataset Penyakit Liver .....	34
Tabel 4. 4 Heatmap Korelasi .....	35
Tabel 4. 5 Feature Selection perhitungan Fisher Score .....	37
Tabel 4. 6 Menampilkan Seleksi Fitur Berdasarkan Fisher Score .....	40
Tabel 4. 7 Menampilkan Teknik SMOTE untuk Menyeimbangkan Data .....	41
Tabel 4. 8 Menampilkan Perbandingan Sampel Sebelum dan Sesudah SMOTE .....	42
Tabel 4. 9 Menampilkan Pembagian Dataset dengan Train-Test Split .....	47
Tabel 4. 10 Jumlah Sampel Data Training dan Data Testing .....	48
Tabel 4. 11 Model Random Forest (Tanpa Optimasi) .....	48
Tabel 4. 12 Confusion Matrix Random Forest Tanpa Optimasi .....	51
Tabel 4. 13 Random Forest Menggunakan Optimasi Model Random Search .....	53
Tabel 4. 14 Confusion Matrix Random Forest Menggunakan Optimasi Model Random Search .....	57
Tabel 4. 15 Perbandingan Hasil Evaluasi Model Sebelum dan Sesudah Optimasi	60

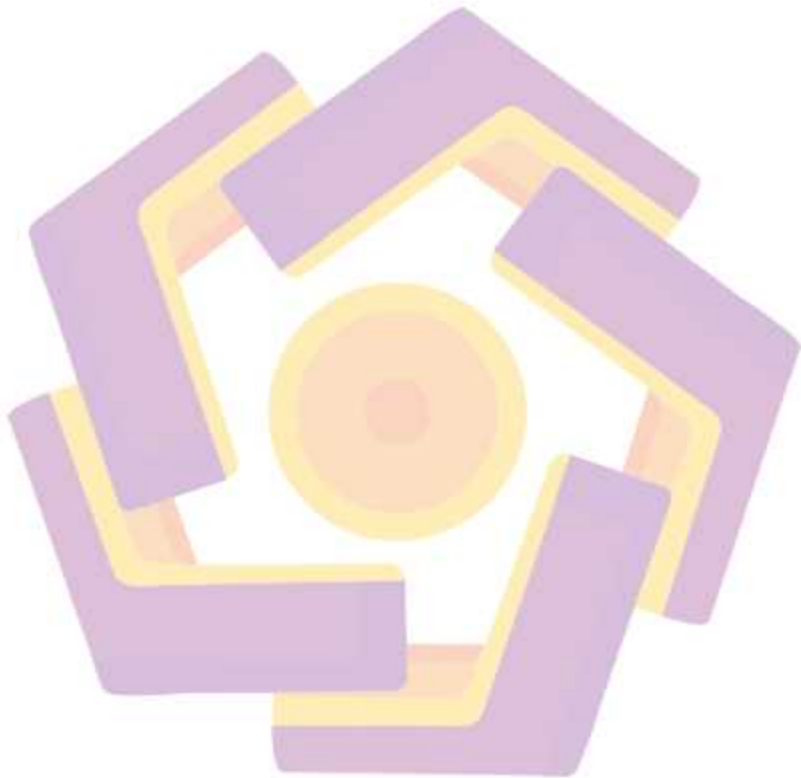
## DAFTAR GAMBAR

Gambar 2. 1 Penjelasan Gambar Dari Hati Sehat Dan Penyakit Hati Lainnya .....	13
Gambar 2. 2 Algoritma Random Forest.....	18
Gambar 3. 1 Alur Penelitian .....	23
Gambar 4. 1 Import Data .....	33
Gambar 4. 2 Informasi Struktur Dataset .....	34
Gambar 4. 3 Statistik Deskriptif Dataset .....	35
Gambar 4. 4 Heatmap Matriks Korelasi Antar Fitur .....	36
Gambar 4. 5 hasil perhitungan Fisher Score .....	39
Gambar 4. 6 pemodelan berdasarkan threshold tetap (Hasil Seleksi Fitur).....	40
Gambar 4. 7 Distribusi Kelas Sebelum SMOTE .....	45
Gambar 4. 8 Distribusi Kelas Setelah SMOTE .....	46
Gambar 4. 9 Hasil X_train dan X_test.....	48
Gambar 4. 10 Hasil Evaluasi Model Random Forest Tanpa Optimasi.....	50
Gambar 4. 11 Confusion Matrix Model Random Forest Tanpa Optimasi .....	52
Gambar 4. 12 Hasil Evaluasi Model Random Forest Setelah Optimasi .....	57
Gambar 4. 13 Confusion Matrix Setelah Optimasi Random Search .....	59
Gambar 4. 14 Hasil Deployment .....	61



## DAFTAR LAMPIRAN

Lampiran 1. Code Github .....	71
Lampiran 2 Hasil Deployment.....	71





## DAFTAR LAMBANG DAN SINGKATAN

$\hat{y}$	adalah hasil prediksi akhir.
$n$	adalah jumlah pohon dalam hutan (forest).
$T_i(x)$	adalah prediksi dari pohon ke- $i$ untuk input $x$
Gini	adalah ukuran seberapa tidak murni atau campur aduk suatu node dalam decision tree.
$\mathcal{K}$	adalah jumlah total kelas dalam dataset.
$p_k$	adalah proporsi atau probabilitas data yang termasuk dalam kelas ke- $k$ di node tersebut.
$(p_k)^2$	adalah jumlah dari kuadrat probabilitas setiap kelas di dalam node
$\theta_i$	adalah nilai yang dipilih untuk hyperparameter ke- $i$ (vektor parameter model)
$P(H_i)$	adalah distribusi probabilitas untuk memilih nilai $\theta_i$ dari domain $H_i$ .
$\sum$	Simbol sigma ( $\Sigma$ ) adalah notasi penjumlahan. Artinya untuk menjumlahkan suatu ekspresi berbagai nilai dari suatu indeks

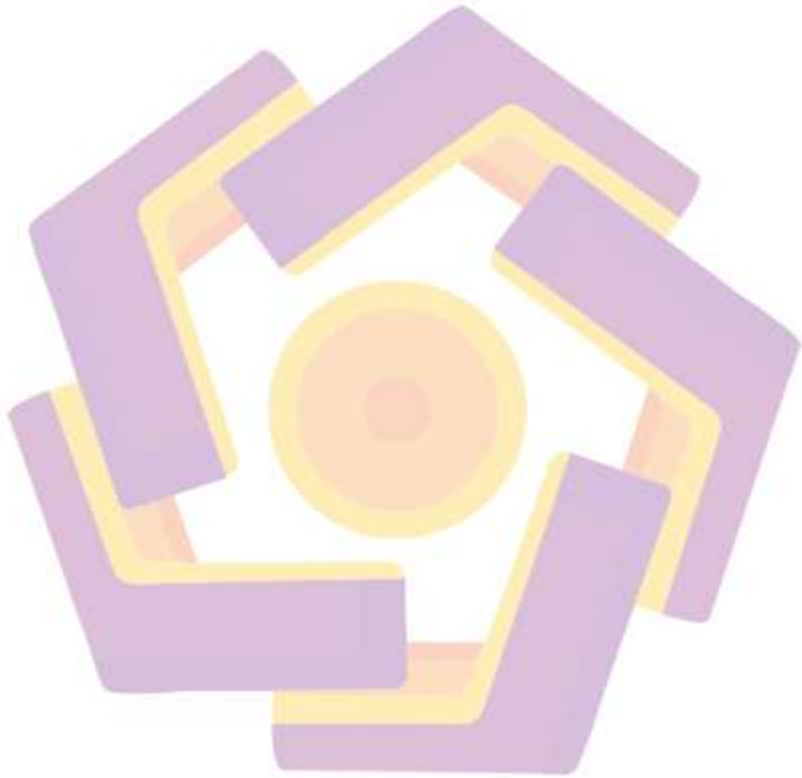
## DAFTAR ISTILAH



Liver	Salah satu penyebab utama morbiditas dan mortalitas global yang menyebabkan berbagai jenis penyakit hati, seperti perlemakan hati (fatty liver), sirosis, hepatitis, kanker hati, dan penyakit hati kronis.
Random Forest	Algoritma machine learning berbasis ensemble yang terdiri dari banyak pohon keputusan untuk meningkatkan akurasi dan mengurangi overfitting.
Akurasi	Metrik evaluasi yang mengukur persentase prediksi yang benar dari keseluruhan data.
Presisi	Metrik yang mengukur ketepatan model dalam memprediksi kelas positif, dihitung sebagai rasio antara True Positive (TP) dan total prediksi positif (TP + FP).
Recall	Metrik yang mengukur seberapa baik model mendeteksi seluruh instance positif, dihitung sebagai rasio antara True Positive (TP) dan total data positif sebenarnya (TP + FN).
F1-Score	Rata-rata harmonis antara presisi dan recall, digunakan untuk menyeimbangkan keduanya dalam evaluasi model.
Optimasi	Proses mencari nilai parameter terbaik untuk meningkatkan kinerja model.
Random Search	Metode optimasi hyperparameter yang memilih kombinasi parameter secara acak untuk menemukan hasil terbaik dengan lebih efisien dibandingkan Grid Search.

Machine Learning

Cabang kecerdasan buatan yang memungkinkan komputer belajar dari data untuk membuat prediksi atau keputusan tanpa pemrograman eksplisit.



## INTISARI

Hati adalah organ vital manusia yang memiliki fungsi kompleks dan beragam. Salah satu penyakit yang menyerang hati adalah hepatitis atau liver. Deteksi dini sangat penting untuk memungkinkan intervensi yang lebih efektif dan memperlambat perkembangan penyakit. Namun, proses diagnosis penyakit Liver seringkali menghadapi tantangan, terutama dalam mendeteksi tahap awal penyakit dari data medis yang kompleks dan beragam. Penelitian ini bertujuan untuk mengoptimalkan algoritma Random Forest menggunakan metode Random Search pada deteksi penyakit Liver. Algoritma Random Forest diterapkan sebagai model utama dalam penelitian ini, sementara optimasi hyperparameter dilakukan menggunakan metode Random Search untuk meningkatkan performa model. Hasil penelitian menampilkan bahwa model Random Forest tanpa optimasi menghasilkan akurasi sebesar 93%. Setelah dilakukan optimasi hyperparameter, akurasi model meningkat menjadi 94%. Kesimpulannya, penerapan optimasi hyperparameter menggunakan metode Random Search berhasil meningkatkan performa model Random Forest. Model yang dihasilkan dapat memberikan hasil prediksi yang lebih akurat.

**Kata kunci:** Liver, Machine Learning, Random Forest, Optimasi, Random Search.

## ABSTRACT

The liver is a vital human organ with complex and diverse functions. One of the diseases that affect the liver is hepatitis or liver disease. Early detection is crucial to enable more effective intervention and slow the progression of the disease. However, diagnosing liver disease often faces challenges, especially in detecting the early stages of the disease from complex and diverse medical data. This study aims to optimize the *Random Forest* algorithm using the *Random Search* method for liver disease detection. The *Random Forest* algorithm is applied as the primary model in this research, while hyperparameter optimization is performed using the *Random Search* method to enhance model performance. The results show that the *Random Forest* model without optimization achieves an accuracy of 93%. After hyperparameter optimization, the model's accuracy increases to 94%. In conclusion, applying hyperparameter optimization using the *Random Search* method successfully improves the performance of the *Random Forest* model. The resulting model provides more accurate predictions.

**Keywords:** Liver, Machine Learning, Random Forest, Optimization, Random Search