

**PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE DAN RANDOM FOREST UNTUK PREDIKSI
PENYAKIT JANTUNG MENGGUNAKAN METODE
FEATURE SELECTION ANNOVA F VALUE**

SKRIPSI SARJANA INFORMATIKA

Oleh:

Andini Putri Humaira

227064516139



**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL**

2026

**PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE DAN RANDOM FOREST UNTUK PREDIKSI
PENYAKIT JANTUNG MENGGUNAKAN METODE
FEATURE SELECTION ANNOVA F VALUE**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer dari Fakultas Teknologi Komunikasi dan Informatika

Oleh:

Andini Putri Humaira

227064516139



**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2026

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

TUGAS AKHIR

**PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE DAN RANDOM FOREST UNTUK PREDIKSI
PENYAKIT JANTUNG MENGGUNAKAN METODE
FEATURE SELECTION ANNOVA F VALUE**



Andini Putri Humaira

227064516139

Dosen Pembimbing 1

A handwritten signature in black ink, appearing to read 'Sari Ningsih'.

(Sari Ningsih, S.Si., M.M.,)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN METODE FEATURE SELECTION ANNOVA F VALUE

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.



Jakarta, 27 Februari 2026



Andini Putri Humaira

227064516139

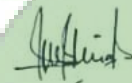
LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

**PERBANDINGAN ALGORITMA SUPPORT VECTOR
MACHINE DAN RANDOM FOREST UNTUK PREDIKSI
PENYAKIT JANTUNG MENGGUNAKAN METODE
FEATURE SELECTION ANNOVA F VALUE**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2025-2026 pada tanggal 25 Februari Tahun 2026


Dosen Pembimbing I



Sari Ningsih, S.Si., M.M.,

NIDN 0302066701

Ketua Program Studi



Rahli Widyadarmasari, ST.,

MM., MMSI

NIDN 0301038302

UNIVERSITAS NASIONAL

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Andini Putri Humaira
 NPM : 227064516139
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Informatika
 Tanggal Sidang : 25 Februari 2026

JUDUL DALAM BAHASA INDONESIA :

**PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE
 DAN RANDOM FOREST UNTUK PREDIKSI PENYAKIT
 JANTUNG MENGGUNAKAN METODE FEATURE SELECTION
 ANNOVA F VALUE**

JUDUL DALAM BAHASA INGGRIS :

**COMPARISON OF SUPPORT VECTOR MACHINE AND
 RANDOM FOREST ALGORITHMS FOR HEART DISEASE
 PREDICTION USING FEATURE SELECTION ANNOVA F VALUE
 METHOD**

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Pembimbing 2	Mahasiswa
TGL : 27 Februari 2026	TGL : 27 Februari 2026	TGL : 27 Februari 2026
 Sari Ningsih, S.si., M.M.,	 Ratih Titi Kaulasari S.T.M.M., M.M.S.I	 Andini Putri Humaira

KATA PENGANTAR

Puji Syukur penulis panjatkan atas kehadiran Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Perbandingan Algoritma Support Vector Machine dan Random Forest Untuk Prediksi Penyakit Jantung Menggunakan Metode Feature Selection Anova F Value”. Adapun tujuan dari penyusunan skripsi ini ialah untuk memenuhi salah satu syarat memperoleh gelar Sarjana Komputer di Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional Jakarta.

Dalam penyusunan skripsi ini penulis banyak menerima bantuan, bimbingan serta dukungan dari berbagai pihak, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Maka dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dalam penyusunan skripsi ini, terutama kepada yang terhormat:

1. Bapak Dr. Agung Triayudi, S.Kom.,M.Kom Selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional
2. Ibu Ir. Endah Tri Esti Handayani, MMSI Selaku Wakil Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional
3. Ibu Ratih Titi Komalasari, ST, MM, MMSI Selaku Ketua Program Studi Informatika Universitas Nasional
4. Ibu Sari Ningsih, S.Si., M.M., selaku Dosen Pembimbing yang telah meluangkan waktu, pikiran, dan tenaga untuk memberikan bimbingan, arahan, serta saran yang sangat berharga dalam penyusunan skripsi ini. Berkat ketelitian, kesabaran, dan dukungan beliau, saya dapat menyelesaikan penelitian ini dengan baik.
5. Para Dosen dan Seluruh Staff akademik Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional yang telah memberikan bekal ilmu yang bermanfaat
6. Kedua orangtua tercinta dan keluarga saya sayangi yang senantiasa mencurahkan segenap kasih sayang yang tiada henti-hentinya, doa, motivasi, nasehat, serta kesabaran yang begitu besar.
7. Terima kasih saya sampaikan kepada sahabat-sahabat saya serta seseorang yang sangat spesial bagi saya, yang selalu hadir memberikan dukungan, bantuan, dan masukan selama proses penyusunan skripsi ini. Kehadiran mereka benar-benar berarti dan menjadi kekuatan bagi saya untuk terus melangkah. Terima kasih juga karena telah bersedia mendengarkan setiap keluh kesah saya selama ini, selalu memberikan

semangat di saat saya merasa lelah, serta tak henti menghibur saya dengan gaya dan caranya sendiri. Semua dukungan itu membuat saya mampu menjalani dan menyelesaikan penelitian ini dengan lebih baik.

8. Terakhir penulis ingin mengucapkan terimakasih kepada seseorang yang mungkin sering terlupakan, ya penulis ingin mengucapkan terimakasih kepada diri sendiri karena tetap bertahan sampai pada tahap ini. Terimakasih karena tidak menyerah dan berani melawan rasa takut, serta keraguan terbesar dalam diri, terimakasih untuk tidak menyerah **tetapi memilih melangkah dan terus berjuang walau kadang tak tau arah dan sering merasa terlambat dari yang lain.** Sampai ditahap ini bukanlah hal yang mudah bagi anak kecil yang perjuangannya **dipenuhi dengan benturan besar,** perjalanan mu sangat tidak mudah tapi kau sanggup **melewati badai itu.** Saya bangga padamu, saya tau perjalanan ini belum selesai, masih banyak ketidakpastian dan luka yang mungkin datang tetapi semoga kamu mampu melewatinya dan ingat kamu pantas untuk bahagia, kamu berhak bermimpi, dan kamu layak untuk sampai di tujuan itu. Teruslah hidup dengan hati yang jujur, dan berjalan dengan niat yang baik. Terimakasih Andini Putri Humaira kamu hebat sudah sejauh ini.

Penulis mengakui bahwa skripsi ini memiliki kekurangan dalam berbagai aspek, termasuk materi, isi, dan teknik penyajian. Hal ini disebabkan oleh keterbatasan pengetahuan dan kemampuan penulis. Oleh karena itu, penulis sangat menghargai kritik dan saran yang bersifat membangun dari semua pihak, dan berharap agar kontribusi tersebut dapat meningkatkan kualitas skripsi ini ke arah yang lebih baik.

ABSTRAK

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia dan terus meningkat setiap tahunnya. Dalam penelitian ini, dilakukan perbandingan performa algoritma Support Vector Machine (SVM) dan Random Forest (RF) dalam memprediksi risiko penyakit jantung berdasarkan dataset yang diperoleh dari Kaggle. Untuk meningkatkan akurasi model, digunakan metode seleksi fitur ANOVA F-value guna memilih fitur yang paling berpengaruh terhadap prediksi penyakit jantung. Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki performa lebih baik dibandingkan SVM, dengan akurasi 99%, sedangkan SVM hanya mencapai 97%. Evaluasi menggunakan confusion matrix juga menunjukkan bahwa Random Forest memiliki tingkat kesalahan klasifikasi yang lebih rendah dibandingkan SVM. Dengan demikian, Random Forest lebih efektif dalam memprediksi risiko penyakit jantung dan dapat digunakan sebagai dasar dalam pengembangan aplikasi deteksi penyakit jantung berbasis machine learning. Penelitian ini diharapkan dapat memberikan kontribusi dalam pemanfaatan teknologi kecerdasan buatan untuk meningkatkan deteksi dini serta mendukung upaya pencegahan penyakit jantung.

Kata Kunci: Penyakit Jantung, Machine Learning, SVM, Random Forest, ANOVA F-value, Feature Selection.



ABSTRACT

Heart disease is one of the leading causes of death worldwide and continues to increase every year. In this study, a comparison of the performance of the Support Vector Machine (SVM) and Random Forest (RF) algorithms was conducted to predict the risk of heart disease based on a dataset obtained from Kaggle. To improve model accuracy, the ANOVA F-value feature selection method was used to select the most influential features for heart disease prediction. The results show that the Random Forest algorithm performs better than SVM, achieving an accuracy of 99%, while SVM reaches only 97%. Evaluation using a confusion matrix also indicates that Random Forest has a lower classification error rate compared to SVM. Therefore, Random Forest is more effective in predicting heart disease risk and can be used as a basis for developing machine-learning-based heart disease detection applications. This study is expected to contribute to the utilization of artificial intelligence technology to enhance early detection and support efforts to prevent heart disease.

Keywords: Heart Disease, Machine Learning, SVM, Random Forest, ANOVA F-value, Feature Selection.



DAFTAR ISI

HALAMAN PENGESAHAN	ii
KATA PENGANTAR	vi
ABSTRAK.....	viii
ABSTRACT.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	2
1.3 Rumusan masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Batasan Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Teori yang Relevan.....	5
2.2 Studi Literatur	6
2.3 Landasan teori.....	15
2.2.1 Penyakit Jantung.....	15
2.2.2 Python.....	15
2.2.3 Streamlit.....	15
2.2.4 Machine Learning	16
2.2.5 Algoritma Suport Vector Machine.....	16
2.2.6 Algoritma Random Forest.....	17
2.2.7 SelectBest Anova F value	18
2.2.8 Confusion matrix	19
BAB III METODE PENELITIAN	21
3.1 Waktu Penelitian	21
3.2 Penentuan Object Penelitian	22
3.3 Fokus penelitian	22
3.4 Sumber Data.....	22

3.5	Teknis Pengumpulan Data	23
3.6	Dataset.....	23
3.7	Desain Penelitian	25
3.7.1	Tahapan Pendahuluan	26
3.7.2	Tahapan Tinjauan Pustaka.....	26
3.7.3	Tahapan Pengumpulan Dataset.....	26
3.7.4	Tahapan Perancangan Model SVM dan RF.....	26
3.7.5	Tahapan Perancangan Streamlit.....	27
3.7.6	Tahapan Hasil Akhir	27
3.8	Metode dan Implementasi Algoritma.....	28
3.8.1	Tahap Preprocessing Data.....	28
3.8.2	Seleksi Fitur Menggunakan ANOVA F-Value	29
3.8.3	Implementasi Support Vector Machine (SVM)	30
3.8.4	Implementasi Random Forest	31
3.8.5	Evaluasi Model	31
BAB IV HASIL DAN PEMBAHASAN		33
4.1	Perancangan Sistem Prediksi Penyakit Jantung Menggunakan Metode ANOVA F-value dan Algoritma SVM serta Random Forest	33
4.2	Perancangan model machine learning menggunakan algoritma Support Vector Machine dan Random Forest dengan anova F value	35
4.2.1	Import Dataset.....	35
4.2.2	Preprocessing dan Split data.....	36
4.2.3	Metode Anova F value	38
4.2.4	Model Machine Learning SVM dan RF.....	40
4.2.5	Confusion Matrix	43
4.2.6	Uji Hipotesis ANOVA F-Value	49
4.3	Perbandingan hasil akurasi menggunakan anova f score dengan akurasi yang tidak menggunakan anova f score.....	50
4.3.1	Analisis Perbandingan Model.....	51
4.4	Tampilan Aplikasi	52
4.4.1	Heart Disease Classification System	52
BAB V KESIMPULAN DAN SARAN.....		57
5.1	Kesimpulan	57
5.2	Saran	57



DAFTAR GAMBAR

Gambar 2. 1 Algoritma Support Vector Machine	17
Gambar 3. 1 Desain Penelitian.....	25
Gambar 4. 1 Rancang Model Machine Learning.....	33
Gambar 4. 2 Source Code Dataset.....	35
Gambar 4. 3 Source Code Preprocessing dan Split Data.....	36
Gambar 4. 4 Source Code Anova F Value	38
Gambar 4. 5 Source Code Model Machine Learning SVM dan Random RF	40
Gambar 4. 6 Hasil Akurasi Algoritma Random Forest.....	41
Gambar 4. 7 Hasil Akurasi Algoritma SVM.....	42
Gambar 4. 8 Source Code Confusion Matrix.....	43
Gambar 4. 9 Hasil Confusion matrix model Algoritma Random Forest	44
Gambar 4. 10 Hasil Confusion matrix Algoritma Support Vector Machine.....	46
Gambar 4. 11 Perbandingan Akurasi Random Forest dan SVM dengan ANOVA F-Value.....	50
Gambar 4. 12 Perbandingan Akurasi Random Forest dan SVM tanpa ANOVA F-Value	51
Gambar 4. 13 Gambar Tampilan Awal Streamlit.....	52
Gambar 4. 14 Dataset Preview dan Pemilihan Target.....	53
Gambar 4. 15 Hasil Seleksi Fitur Menggunakan ANOVA	54
Gambar 4. 16 Visualisasi Top 10 Feature Menggunakan ANOVA F-Value	55
Gambar 4. 17 Perbandingan Hasil Evaluasi Model Random Forest dan SVM.....	56



DAFTAR TABEL

Tabel 2. 1 Studi Literatur	6
Tabel 2. 2 Confusion matrix.....	19
Tabel 3. 1 Timeline Penelitian	21
Tabel 3. 2 Variable Independen.....	24
Tabel 3. 3 Variable Dependen	24
Tabel 4. 2 Socre F All Variable	39
Tabel 4. 3 Variable Selection Form Anova F Value	40
Tabel 4. 4 Hasil Confusion Matrix Algoritma Random Forest.....	44
Tabel 4. 5 Hasil Confusion matrix model Algoritma Support Vector Machine.....	47

