

Prediksi Risiko Serangan Jantung Menggunakan Algoritma Random Forest Berdasarkan Faktor Klinis dan Riwayat Medis Pasien

Arkananta Hamzah Firjatullah¹, Muhammad Vhijar Rainaldany², Ryan Firas Anderasta³ Yamin Nuryamin⁴, Ade Priyatna⁵

^{1,2,3} Mahasiswa, Teknologi Informasi Universitas Bina Sarana Informatika,

^{4,5} Dosen, Teknologi Informasi Universitas Bina Sarana Informatika.

¹arzah7569@gmail.com, ²vhijardanny2@gmail.com, ³ryanfirasanderasta@gmail.com, ⁴yamin.yny@bsi.ac.id, ⁵ade.aeq@bsi.ac.id

Article Info

Article history:

Received October 2, 2025

Accepted November 10, 2025

Published January 3, 2026

Kata Kunci:

Prediksi Risiko Serangan Jantung

Algoritma Random Forest

Faktor Klinis

Riwayat Medis Pasien

Akurasi Prediksi

ABSTRAK

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia, sehingga deteksi dini sangat penting untuk membantu proses diagnosis klinis. Penelitian ini bertujuan memprediksi risiko serangan jantung menggunakan algoritma Random Forest berdasarkan 14 parameter klinis, seperti usia, tekanan darah, kolesterol, detak jantung maksimum, oldpeak, dan tipe nyeri dada. Dataset berjumlah 1.025 data pasien yang telah melalui proses pembersihan, encoding variabel kategorikal, dan standardisasi fitur. Model dilatih menggunakan pembagian data 80% pelatihan dan 20% pengujian. Hasil evaluasi menunjukkan akurasi sebesar 98,53%, precision dan recall $\geq 0,97$, serta performa prediksi yang sangat stabil berdasarkan Confusion Matrix. Analisis feature importance mengungkapkan bahwa chest pain type, oldpeak, dan Max heart rate merupakan faktor yang paling berpengaruh. Model Random Forest terbukti efektif sebagai alat pendukung keputusan dalam identifikasi risiko serangan jantung.



Corresponding Author:

Arkananta Hamzah Firjatullah

Teknologi Informasi Universitas Bina Sarana Informatika

Email: *arzah7569@gmail.com

1. PENDAHULUAN

Penyakit jantung masih menjadi salah satu penyebab kematian terbesar secara. Deteksi dini terhadap risiko serangan jantung sangat penting untuk mencegah komplikasi serius dan meningkatkan peluang keselamatan pasien. Jantung merupakan salah satu bagian dari organ vital manusia yang perlu dijaga untuk keberlangsungan hidup manusia. Salah satu hal dalam dunia kesehatan yang saat ini menjadi sorotan adalah berbagai penyakit yang dapat dialami oleh organ vital manusia termasuk jantung (Elektronik et al., 2023). Di Indonesia, penyakit jantung juga menjadi penyebab utama kematian, yang diperburuk oleh keterbatasan akses ke fasilitas kesehatan, terutama di daerah terpencil. Metode diagnostik tradisional, seperti pemeriksaan fisik dan EKG, sering kali tidak cukup akurat dalam memprediksi serangan jantung (Mandias & Manoppo, 2025). sekitar 17,9 juta kematian setiap tahun. Diagnosis dini dan akurat sangat penting untuk pengobatan yang efektif, namun ketidakseimbangan kelas dalam dataset medis sering menyebabkan bias pada model prediktif, khususnya dalam mengidentifikasi pasien dengan penyakit jantung (kelas minoritas) (Pranajaya & Susanto, 2025). Bahaya penyakit ini meliputi serangan jantung mendadak, gagal jantung, hingga stroke akibat komplikasi pembuluh darah (Rahmada et al., 2024). Deteksi dini terhadap risiko penyakit jantung menjadi krusial dalam menurunkan angka kematian dan meningkatkan kualitas hidup pasien (Saputra & Sugihartono,

2025).

Perkembangan teknologi machine learning telah membuka era baru dalam prediksi medis, terutama untuk penyakit kardiovaskular yang memerlukan analisis multivariabel kompleks. Berbagai algoritma seperti Support Vector Machine (SVM), Naïve Bayes, dan Neural Network telah diuji untuk mendeteksi risiko serangan jantung, namun sering kali terkendala oleh ketidakseimbangan dataset medis dan overfitting pada fitur klinis yang bersifat heterogeny (Awaludin & Yasin, 2020). Penelitian terkini menunjukkan bahwa pendekatan ensemble learning, khususnya Random Forest, mampu mengatasi tantangan ini dengan menggabungkan beberapa pohon keputusan untuk menghasilkan prediksi yang lebih stabil dan akurat (Lamine et al., 2023). Keunggulan ini krusial dalam konteks kesehatan, di mana kesalahan prediksi (false negative/false positive) dapat berdampak fatal pada keselamatan pasien. Oleh karena itu, pengembangan model prediksi berbasis Random Forest menjadi solusi strategis untuk mengoptimalkan deteksi dini penyakit jantung.

Implementasi Random Forest dalam prediksi risiko jantung juga memberikan nilai tambah melalui analisis *feature importance* yang transparan dan dapat diinterpretasikan secara klinis. Berbeda dengan model "black box" seperti Deep Learning, Random Forest memungkinkan identifikasi faktor dominan yang memengaruhi prediksi seperti tipe nyeri dada (*chest pain type*), depresi ST (*oldpeak*), dan detak jantung maksimum sehingga hasilnya dapat divalidasi oleh tenaga medis. Hal ini sejalan dengan kebutuhan praktik klinis modern yang mengedepankan *explainable AI* (XAI) untuk mendukung keputusan diagnostic (Awaludin & Mantik, 2023). Selain itu, kemampuan algoritma ini menangani data kategorikal dan numerik tanpa preprocessing ekstensif menjadikannya efisien untuk diintegrasikan ke dalam sistem pendukung keputusan (clinical decision support system) di fasilitas kesehatan, termasuk di daerah terpencil dengan keterbatasan sumber daya. Algoritma Random Forest merupakan metode klasifikasi berbasis ensemble yang mampu menangani data multivariabel dan menghasilkan performa prediksi stabil (Awaludin, 2018). Penelitian ini menerapkan Random Forest untuk memprediksi risiko serangan jantung berdasarkan 14 fitur klinis seperti usia, tekanan darah, kolesterol, detak jantung maksimum, dan tipe nyeri dada. Tujuannya adalah mengevaluasi performa model dan mengidentifikasi faktor klinis paling berpengaruh terhadap risiko serangan jantung.

Dengan berkembangnya data klinis dan pemanfaatan machine learning, proses prediksi risiko penyakit kini dapat dilakukan dengan lebih cepat dan akurat. Algoritma Random Forest merupakan metode klasifikasi berbasis ensemble yang mampu menangani data multivariabel dan menghasilkan performa prediksi stabil. Penelitian ini menerapkan Random Forest untuk memprediksi risiko serangan jantung berdasarkan 14 fitur klinis seperti usia, tekanan darah, kolesterol, detak jantung maksimum, dan tipe nyeri dada. Tujuannya adalah mengevaluasi performa model dan mengidentifikasi faktor klinis paling berpengaruh terhadap risiko serangan jantung.

2. METODE

State of the art Penelitian

No	Nama Peneliti (Tahun)	Pembahasan (Masalah dan Solusi)	Hasil (Kuantitatif)
1	Elektronik et al. (2023)	Masalah: Keterbatasan akurasi prediksi penyakit jantung koroner. Solusi: Penerapan algoritma Random Forest untuk analisis faktor risiko.	Tidak disebutkan (fokus pada analisis algoritma)
2	Mandias & Manoppo (2025)	Masalah: Metode diagnostik tradisional (pemeriksaan fisik/EKG) kurang akurat. Solusi: Penggunaan model machine learning untuk prediksi dini.	Tidak disebutkan (penelitian konseptual)
3	Pranajaya & Susanto (2025)	Masalah: Ketidakseimbangan kelas dataset medis menyebabkan bias prediksi. Solusi: Optimasi Random Forest dengan SMOTEENN (penanganan data tidak seimbang) dan Grid Search.	Akurasi: Tidak disebutkan Fokus: Peningkatan performa model pada data minoritas

No	Nama Peneliti (Tahun)	Pembahasan (Masalah dan Solusi)	Hasil (Kuantitatif)
4	Rahmada et al. (2024)	Masalah: Akurasi prediksi penyakit jantung masih rendah. Solusi: Kombinasi SMOTEENN dengan Random Forest untuk penanganan data tidak seimbang.	Akurasi: Tidak disebutkan Recall: Peningkatan signifikan pada kelas minoritas
5	Saputra & Sugihartono (2025)	Masalah: Keterbatasan model tradisional dalam memprediksi risiko jantung. Solusi: Evaluasi model LSTM (Long Short-Term Memory) untuk prediksi berbasis time-series.	Akurasi: Tidak disebutkan Fokus: Perbandingan performa LSTM vs model lain
6	Penelitian Saat Ini (2020)	Masalah: Kebutuhan deteksi dini risiko serangan jantung dengan akurasi tinggi. Solusi: Random Forest dengan 14 fitur klinis + preprocessing data (encoding, standarisasi).	Akurasi: 98,53% Precision/Recall: $\geq 0,97$ F1-Score: Mendekati 1

Penelitian ini menggunakan metode kuantitatif karena seluruh variabel penelitian berbentuk data numerik dan kategorikal yang dapat diukur secara objektif. Pendekatan kuantitatif memungkinkan peneliti melakukan pengukuran, pengujian, dan evaluasi performa model secara terstruktur melalui perhitungan statistik dan metrik evaluasi. Berikut gambar alur penelitian

2.1 Pengumpulan Data

Dataset diperoleh dari sumber Kaggle yang berisi data medis terkait faktor risiko penyakit jantung. Dataset tersebut memiliki 14 variabel, seperti umur, jenis kelamin, tekanan darah, kolesterol, tingkat ST depression (oldpeak), thalassemia, serta variabel target (0 = tidak berisiko, 1 = berisiko). Tahap ini sangat penting karena kualitas dataset akan menentukan kualitas model prediksi yang dihasilkan. Dataset dapat diperoleh melalui :

<https://www.kaggle.com/datasets/ketangangal/heart-disease-dataset-uci/data>

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	age	sex	chest_pain	resting_blood_pressure	cholesterol	fasting_blood_sugar	resting_ectg	max_heart_rate	exercise_induced_angina	oldpeak	slope	vessels	ct	thalassemia	target
2	52	Male	Typical	an	125	212	Lower than ST-T wave	168	No	1.0	Downsloping	Two	Reversible	0	
3	53	Male	Typical	an	140	203	Greater than Normal	155	Yes	3.1	Upsloping	Zero	Reversible	0	
4	70	Male	Typical	an	145	174	Lower than ST-T wave	125	Yes	2.6	Upsloping	Zero	Reversible	0	
5	61	Male	Typical	an	148	203	Lower than ST-T wave	161	No	0.0	Downsloping	One	Reversible	0	
6	62	Female	Typical	an	138	294	Greater than ST-T wave	106	No	1.9	Flat	Three	Fixed Defect	0	
7	58	Female	Typical	an	100	248	Lower than Normal	122	No	1.0	Flat	Zero	Fixed Defect	1	
8	58	Male	Typical	an	114	318	Lower than Left ventricular hypertrophy	140	No	4.4	Upsloping	Three	Normal	0	
9	55	Male	Typical	an	160	289	Lower than Normal	145	Yes	0.8	Flat	One	Reversible	0	
10	46	Male	Typical	an	120	249	Lower than Normal	144	No	0.8	Downsloping	Zero	Reversible	0	
11	54	Male	Typical	an	122	286	Lower than Normal	116	Yes	3.2	Flat	Two	Fixed Defect	0	
12	71	Female	Typical	an	112	149	Lower than ST-T wave	125	No	1.6	Flat	Zero	Fixed Defect	1	
13	43	Female	Typical	an	132	341	Greater than Normal	136	Yes	3.0	Flat	Zero	Reversible	0	
14	34	Female	Atypical	an	118	210	Lower than ST-T wave	192	No	0.7	Downsloping	Zero	Fixed Defect	1	
15	51	Male	Typical	an	140	298	Lower than ST-T wave	122	Yes	4.2	Flat	Three	Reversible	0	
16	52	Male	Typical	an	128	204	Greater than ST-T wave	156	Yes	1.0	Flat	Zero	No	0	
17	34	Female	Atypical	an	118	210	Lower than ST-T wave	192	No	0.7	Downsloping	Zero	Fixed Defect	1	
18	51	Female	Non-anginal		140	308	Lower than Normal	142	No	1.5	Downsloping	One	Fixed Defect	1	
19	54	Male	Typical	an	124	266	Lower than Normal	109	Yes	2.2	Flat	One	Reversible	0	
20	50	Female	Atypical	an	120	244	Lower than ST-T wave	162	No	1.1	Downsloping	Zero	Fixed Defect	1	
21	58	Male	Non-anginal		140	211	Greater than Normal	165	No	0.0	Downsloping	Zero	Fixed Defect	1	
22	60	Male	Non-anginal		140	185	Lower than Normal	155	No	3.0	Flat	Zero	Fixed Defect	0	
23	67	Female	Typical	an	106	223	Lower than ST-T wave	142	No	0.3	Downsloping	Two	Fixed Defect	1	
24	45	Male	Typical	an	104	208	Lower than Normal	148	Yes	3.0	Flat	Zero	Fixed Defect	1	

Gambar 2 : Dataset Penelitian

2.2 Eksplorasi Data (EDA)

Exploratory Data Analysis dilakukan untuk memahami karakteristik dataset. Pada tahap ini digunakan fungsi seperti `df.info()`, `df.describe()`, dan visualisasi sederhana untuk melihat:

- tipe data (numerik / kategorikal)

- b. jumlah data
- c. nilai yang tidak konsisten.
- d. distribusi variabel target.
- e. potensi outlier.

EDA membantu peneliti mengenali pola awal dan menyiapkan langkah preprocessing yang sesuai.

2.3 Preprocessing Data

Tahap preprocessing dilakukan agar data siap digunakan oleh model machine learning. Proses ini meliputi:

- a. Encoding variabel kategorikal menggunakan One-Hot Encoding.
- b. Konversi tipe data agar model dapat memproses seluruh nilai.
- c. Normalisasi / Standarisasi jika diperlukan.
- d. Memisahkan fitur dan target (X dan y).

Preprocessing yang baik memastikan model menerima data dalam format yang sesuai dan mengurangi potensi bias.

2.4 Pemisahan Data (Train-Test Split)

Dataset dibagi menjadi dua bagian:

- a. Training set (80%) untuk melatih model Random Forest.
- b. Testing set (20%) untuk menguji performa model pada data yang belum pernah dilihat.

Pembagian ini bertujuan mengevaluasi kemampuan model dalam melakukan generalisasi pada data baru.

2.5 Pelatihan Model Random Forest

Pada tahap ini, algoritma Random Forest dilatih menggunakan data training. Random Forest dipilih karena:

- a. mampu menangani data numerik dan kategorikal.
- b. bekerja baik pada dataset dengan banyak fitur.
- c. mengurangi risiko overfitting melalui teknik bagging.
- d. memberikan analisis feature importance yang berguna untuk analisis medis.

Model mempelajari pola hubungan antar fitur untuk memprediksi apakah pasien berisiko mengalami serangan jantung.

2.6 Evaluasi Model

Model yang telah dilatih diuji menggunakan data testing. Evaluasi dilakukan menggunakan:

- a. Akurasi.
- b. Precision.
- c. Recall.
- d. F1-score.
- e. Confusion Matrix

3. Pembahasan

3.1 Hasil Eksplorasi Data

Tahap eksplorasi data dilakukan untuk memahami karakteristik awal dataset yang digunakan. Dataset berisi 1.025 data pasien dengan 14 variabel, terdiri dari 5 fitur numerik dan 8 fitur kategorikal. Struktur data menunjukkan bahwa tidak terdapat nilai kosong (missing values) sehingga seluruh data dapat langsung diproses. Variabel target terdiri dari dua kelas, yaitu 0 (tidak berisiko serangan jantung) dan 1 (berisiko serangan jantung). Struktur dataset memperlihatkan bahwa data memiliki variasi yang baik pada variabel klinis seperti tekanan darah, kolesterol, detak jantung maksimum, serta indikator kondisi jantung lainnya. Informasi ini penting sebagai dasar untuk membangun model prediksi yang mampu mendeteksi pola terkait risiko serangan jantung.

```

*** <class 'pandas.core.frame.DataFrame'>
RangeIndex: 1025 entries, 0 to 1024
Data columns (total 14 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   age                                  1025 non-null   int64
 1   sex                                  1025 non-null   object
 2   chest_pain_type                      1025 non-null   object
 3   resting_blood_pressure              1025 non-null   int64
 4   cholestoral                         1025 non-null   int64
 5   fasting_blood_sugar                 1025 non-null   object
 6   rest_ecg                           1025 non-null   object
 7   Max_heart_rate                     1025 non-null   int64
 8   exercise_induced_angina             1025 non-null   object
 9   oldpeak                             1025 non-null   float64
10   slope                               1025 non-null   object
11   vessels_colored_by_flourosopy       1025 non-null   object
12   thalassemia                         1025 non-null   object
13   target                              1025 non-null   int64
dtypes: float64(1), int64(5), object(8)
memory usage: 112.2+ KB

```



Gambar 3 Hasil Explorasi Data

3.2 Hasil Preprocessing Data

Proses preprocessing dilakukan untuk mempersiapkan data sebelum digunakan dalam pelatihan model. Tahapan ini meliputi pemisahan fitur dan label, serta proses one-hot encoding pada variabel kategorikal seperti jenis kelamin, tipe nyeri dada, hasil elektrokardiogram, dan thalassemia. Encoding diperlukan agar seluruh fitur dapat dipahami oleh algoritma Random Forest yang hanya menerima input numerik. Selain itu, data dibagi menjadi dua bagian menggunakan metode train-test split, yaitu 80% data untuk pelatihan model dan 20% untuk pengujian. Pembagian ini bertujuan menguji kemampuan generalisasi model pada data baru yang belum pernah dilihat sebelumnya.


```

▶ # Preprocessing Data
df = df.dropna()
df = pd.get_dummies(df, drop_first=True)
X = df.drop('target', axis=1)
y = df['target']
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Split Data Train-Test
X_train, X_test, y_train, y_test = train_test_split(
    X_scaled, y, test_size=0.2, random_state=42
)

```

Gambar 4 Preprocessing Data

3.4 Hasil Pelatihan Model Random Forest

Model Random Forest dilatih menggunakan parameter dasar dengan jumlah pohon keputusan (*n_estimators*) sebanyak 100. Algoritma ini dipilih karena memiliki kemampuan yang baik dalam menangani data dengan banyak fitur, mampu mengurangi risiko overfitting, serta menghasilkan nilai feature importance yang membantu dalam interpretasi medis. Model dilatih menggunakan data training dan menghasilkan performa yang sangat baik ketika diuji pada data testing.

```

▶ # Training Model Random Forest
rf = RandomForestClassifier(
    n_estimators=100,
    max_depth=None,
    random_state=42
)

# Melatih Model
rf.fit(X_train, y_train)

# Prediksi
y_pred = rf.predict(X_test)

```

Gambar 5 Hasil Pelatihan Random Forest

3.5 Evaluasi Kinerja Model

Evaluasi dilakukan menggunakan beberapa metrik, yaitu akurasi, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa model mampu mengklasifikasikan risiko serangan jantung dengan sangat baik, dengan nilai akurasi 0.985 atau 98.5%. Nilai precision dan recall untuk kedua kelas mencapai 0.97–1.00, yang menandakan bahwa model mampu mendeteksi pasien berisiko dan tidak berisiko secara seimbang dan akurat. Tabel classification report menunjukkan bahwa model dapat mengidentifikasi kelas 1 (berisiko) dengan sangat baik tanpa menghasilkan banyak kesalahan prediksi. Nilai F1-score yang mendekati 1 pada kedua kelas menunjukkan stabilitas model dalam proses klasifikasi.

... Akurasi Model: 0.9853658536585366					
	precision	recall	f1-score	support	
0	0.97	1.00	0.99	102	
1	1.00	0.97	0.99	103	
accuracy			0.99	205	
macro avg	0.99	0.99	0.99	205	
weighted avg	0.99	0.99	0.99	205	

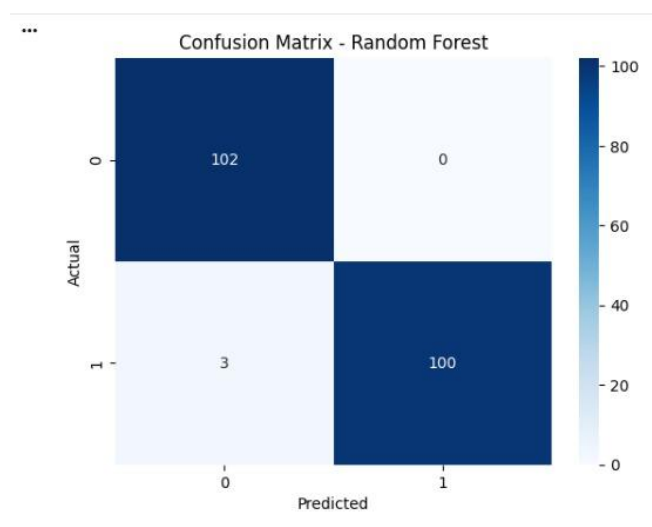
Gambar 6 Evaluasi Kinerja Model

3.6 Analisis Confusion Matrix

Confusion matrix digunakan untuk melihat jumlah prediksi benar dan salah secara lebih rinci. Berdasarkan hasil confusion matrix, model menghasilkan:

- True Negative (TN) = 102
- False Positive (FP) = 0
- False Negative (FN) = 3
- True Positive (TP) = 100

Model tidak menghasilkan false positive, yang berarti tidak ada pasien sehat yang diprediksi salah sebagai pasien berisiko. Namun terdapat tiga kasus di mana model gagal mendeteksi pasien berisiko (FN). Meski demikian, tingkat kesalahan tersebut sangat kecil dibandingkan jumlah data yang diuji. Hasil ini mengindikasikan bahwa model Random Forest mampu memberikan prediksi yang sangat akurat dan dapat diandalkan sebagai sistem pendukung keputusan dalam mendeteksi risiko serangan jantung.



Gambar 7 : Analisis Confusion Matrix

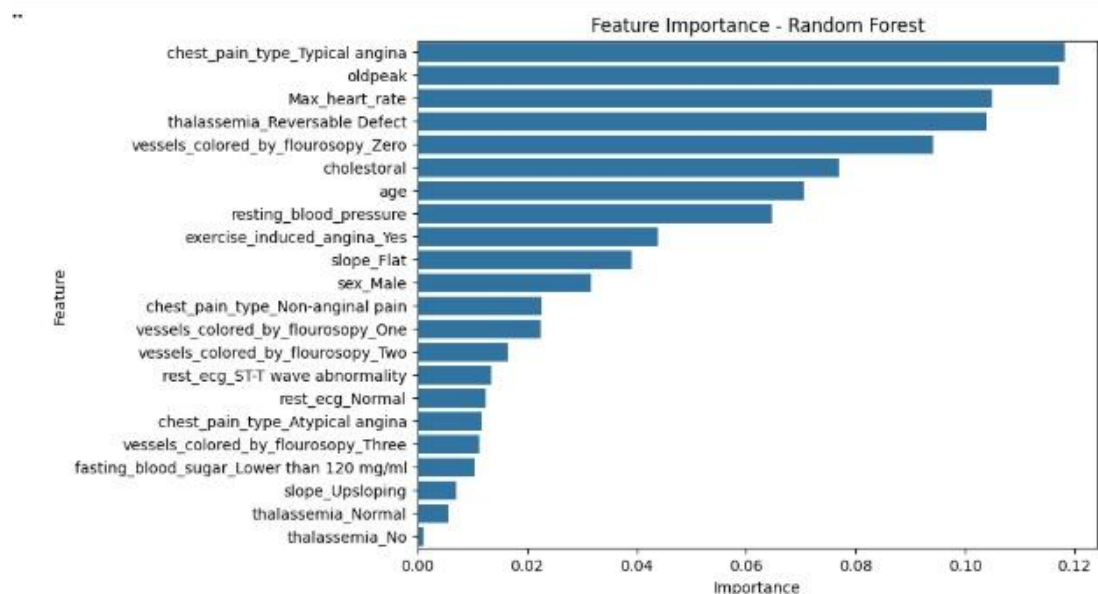
3.7 Analisis feature importance

Analisis feature importance dilakukan untuk mengetahui fitur klinis mana yang paling berpengaruh terhadap prediksi risiko serangan jantung. Berdasarkan hasil model, fitur dengan kontribusi tertinggi adalah:

chest_pain_type_Typical angina
oldpeak
Max_heart_rate

thalassemia_Reversible Defect
vessels_colored_by_flourosopy_Zero
cholestoral
age

Fitur chest pain type dan oldpeak menjadi indikator terbesar, yang secara klinis merupakan gejala umum yang sering digunakan dokter dalam diagnosis penyakit jantung. Sementara itu, max heart rate dan thalassemia reversible defect menunjukkan kontribusi signifikan dalam mengidentifikasi kondisi fungsi jantung.



Gambar 8 Fitur feature importance

Kesimpulan

Penelitian ini menunjukkan bahwa algoritma Random Forest mampu memprediksi risiko serangan jantung dengan sangat baik, ditunjukkan oleh akurasi tinggi sebesar 98,5% serta nilai precision, recall, dan F1-score yang konsisten. Confusion matrix memperlihatkan tingkat kesalahan prediksi yang sangat rendah, sehingga model dapat diandalkan dalam membedakan pasien berisiko dan tidak berisiko. Analisis feature importance mengungkap bahwa variabel jenis nyeri dada, oldpeak, detak jantung maksimum, dan thalassemia merupakan faktor klinis yang paling berpengaruh terhadap risiko serangan jantung. Dengan demikian, Random Forest terbukti efektif sebagai metode prediksi dan memiliki potensi untuk digunakan sebagai sistem pendukung keputusan dalam deteksi dini penyakit jantung.

Daftar Pustaka

- Awaludin, M. (2018). Penerapan Algoritma Rc4 Pada Operasi Xor Untuk Keamanan Pesan Pada Smartphone Berbasis Web. *Jurnal Sistem Informasi Universitas Suryadarma*, 4(1), 16–22. <https://doi.org/10.35968/jsi.v4i1.71>
- Awaludin, M., & Mantik, H. (2023). Penerapan Metode Servqual Pada Skala Likert Untuk Mendapatkan Kualitas Pelayanan Kepuasan Pelanggan. *Jurnal Sistem Informasi Univesitas Suryadarma*, 10(1).
- Awaludin, M., & Yasin, V. (2020). Application Of Oriented Fast And Rotated Brief (Orb) And Bruteforce Hamming In Library Opencv For Classification. *Journal of Information System, Applied, Management, Accounting, and Reserarch*, 4(3), 51–59.
- Elektronik, J., Udayana, I. K., Getrudis, S., Sadipun, I., Anom, G. N., Putra, C., Cs, M., Raya, J., & Unud, K. (2023). Analisis Algoritma Random Forest Dalam Memprediksi Penyakit Jantung Koroner. *Jurnal Elektronik Ilmu Komputer Udayana*, 11(4), 2654–5101.

- Lamine, W., Sarker, J. H., Ghanmi, N., & Biglari, S. (2023). AI-Powered Machine Learning Regulated Emotions in Entrepreneurial Pitching on Investors' Funding Decision. *Colloquium in Information Science and Technology*, 193–198. <https://doi.org/10.1109/CIST56084.2023.10409972>
- Mandias, G. F., & Manoppo, I. J. (2025). *Penerapan Model Machine Learning untuk Memprediksi Serangan Jantung Dini*. 27(2), 193–201.
- Pranajaya, A. E., & Susanto, E. R. (2025). Optimasi Random Forest untuk Prediksi Penyakit Jantung Menggunakan SMOTEENN dan Grid Search Sistem Informasi , Fakultas Teknik dan Ilmu Komputer , Universitas Teknokrat Indonesia , Indonesia Random Forest Optimization for Heart Disease Prediction Using SM. *Jpti*, 5(7), 1965–1979.
- Rahmada, A., Susanto, E. R., Informasi, S., Teknik, F., & Indonesia, U. T. (2024). *Peningkatan Akurasi Prediksi Penyakit Jantung dengan Teknik SMOTEENN pada Algoritma Random Forest Improving Heart Disease Prediction Accuracy with SMOTEENN Technique on Random Forest Algorithm*. 4(12), 795–803.
- Saputra, M. A., & Sugihartono, T. (2025). *Evaluasi Kinerja Model LSTM Untuk Prediksi Risiko Penyakit Jantung Menggunakan Dataset Program Studi Teknik Informasi , Fakultas Teknologi Informasi , ISB Atma Luhur , Indonesia Performance Evaluation of the LSTM Model for Heart Disease Risk Prediction Us*. 5(7), 1823–1833.