

# BI ICT 2025 - Jantung Koroner - S.doc

*by turnitin student*

---

**Submission date:** 23-Jun-2025 11:32PM (UTC-0500)

**Submission ID:** 2704784068

**File name:** BI\_ICT\_2025\_-\_Jantung\_Koroner\_-\_S.doc (172K)

**Word count:** 1484

**Character count:** 9812

## Penerapan Algoritma Machine Learning untuk Diagnosis Penyakit Jantung Koroner

Muhammad Sava Akbar Bastaman, Muhammad Rasyad Andhika Yogaswara

Teknik Informatika; Universitas Kristen Maranatha; Jl. Puri Cipageran Indah 1 Blok G No 36,  
Cipageran, Kec. Cimahi Utara, Kabupaten Bandung, 087731666512; e-mail:  
[2372036@maranatha.ac.id](mailto:2372036@maranatha.ac.id)

Teknik Informatika; Universitas Kristen Maranatha; Jl. Jaka Santika No 14, Kujangsari,  
Kecamatan Bandung Kidul, Kota Bandung, 082262494077; e-mail:  
[2372042@maranatha.ac.id](mailto:2372042@maranatha.ac.id)

\* Korespondensi: e-mail: [xxxx@xxxx.xxx](mailto:xxxx@xxxx.xxx)

Diterima: ...; Review: ...; Disetujui: ...

Cara sitasi: Penulis Pertama, Penulis Kedua, Penulis Ketiga. 2021. Judul Artikel Ilmiah. Bina  
Insani ICT Journal. Vol (no): halaman.

**Abstrak:** Penyakit jantung koroner (PJK) merupakan penyebab utama kematian di dunia. Deteksi dini penyakit ini menjadi penting untuk meningkatkan kualitas hidup pasien. Penelitian ini bertujuan untuk mengembangkan model prediksi diagnosis PJK menggunakan pendekatan machine learning berbasis algoritma Random Forest. Dataset yang digunakan adalah Heart Disease Dataset yang bersumber dari UCI Machine Learning Repository. Data dianalisis menggunakan 13 atribut klinis, dan dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Hasil evaluasi menunjukkan bahwa model Random Forest mampu mencapai akurasi sebesar 88,52% dengan nilai presisi dan recall yang seimbang. Penelitian ini membuktikan bahwa algoritma machine learning, khususnya Random Forest, efektif dalam membantu diagnosis penyakit jantung koroner dan berpotensi untuk diimplementasikan dalam sistem pendukung keputusan medis.

### 1. Pendahuluan

Penyakit jantung koroner (PJK) merupakan salah satu penyakit tidak menular yang menjadi penyebab utama kematian di dunia, termasuk di Indonesia. PJK terjadi ketika aliran darah menuju otot jantung mengalami hambatan akibat penumpukan plak aterosklerotik pada pembuluh darah koroner. Kondisi ini dapat memicu terjadinya serangan jantung yang bersifat fatal apabila tidak terdeteksi dan ditangani secara dini dan tepat. Oleh karena itu, upaya deteksi dini sangat penting dalam rangka pencegahan serta penanganan lebih lanjut terhadap PJK [1].

Seiring dengan perkembangan teknologi, khususnya di bidang kesehatan digital, pemanfaatan machine learning mulai mendapatkan perhatian sebagai pendekatan alternatif dalam proses diagnosis penyakit, termasuk PJK. Machine learning merupakan cabang dari kecerdasan buatan (artificial intelligence) yang memungkinkan sistem komputer mempelajari data dan mengenali pola secara otomatis. Dalam dunia medis, pendekatan ini mampu mengolah data kesehatan seperti tekanan darah, kadar kolesterol, usia, dan riwayat medis pasien untuk memprediksi risiko penyakit jantung dengan tingkat akurasi yang tinggi [2].

Penelitian ini bertujuan untuk menerapkan dan mengevaluasi beberapa algoritma machine learning dalam membantu diagnosis penyakit jantung koroner. Dengan pendekatan ini, diharapkan dapat dikembangkan suatu sistem pendukung keputusan medis yang efisien, akurat, dan dapat digunakan oleh tenaga kesehatan sebagai alat bantu diagnosis. Teknologi ini juga diharapkan mampu memberikan kontribusi terhadap penurunan angka kejadian PJK di masyarakat melalui strategi pencegahan yang lebih tepat sasaran dan berbasis data.

Penelitian ini memiliki kebaruan dalam hal penerapan dan optimalisasi algoritma decision tree, random forest, dan k-nearest neighbors dalam proses klasifikasi dan prediksi risiko penyakit jantung. Dengan membandingkan performa dari masing-masing algoritma berdasarkan parameter evaluasi seperti akurasi, presisi, dan recall, penelitian ini bertujuan untuk memberikan kontribusi terhadap pengembangan sistem prediksi medis yang unggul dan adaptif. Penelitian ini juga menambahkan nilai inovatif melalui integrasi teknologi machine learning dalam pelayanan kesehatan berbasis data, sebagaimana dibahas secara rinci oleh Nareza [1] dan dijelaskan lebih teknis dalam Bab 4 oleh Eduonix Learning Solutions [2].

## 2. Metode Penelitian

Penelitian ini bertujuan untuk membangun sistem klasifikasi penyakit jantung dengan membandingkan tiga algoritma machine learning, yaitu Decision Tree, Random Forest, dan K-Nearest Neighbors (KNN). Tahapan penelitian dilakukan secara sistematis, dimulai dari akuisisi data, pra-pemrosesan, pemodelan, hingga evaluasi kinerja model. Setiap tahapan dijelaskan sebagai berikut:

### Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimental, di mana ketiga algoritma klasifikasi dibandingkan performanya pada dataset yang sama. Penelitian dilakukan secara insiliko, yaitu seluruh proses dilakukan melalui simulasi komputer menggunakan bahasa pemrograman Python dan pustaka pendukung seperti scikit-learn, pandas, dan matplotlib.

### Akuisisi dan Pra-pemrosesan Data

Dataset yang digunakan adalah Cleveland Heart Disease Dataset yang tersedia di UCI Machine Learning Repository [6]. Dataset ini memiliki 303 entri pasien dengan 14 atribut, di antaranya: usia, jenis kelamin, tekanan darah, kolesterol, dan hasil EKG. Label target adalah keberadaan penyakit jantung, yang kemudian dibagi menjadi dua kelas:

Kelas 0: Tidak memiliki penyakit jantung.

Kelas 1: Memiliki indikasi penyakit jantung.

Tahapan pra-pemrosesan meliputi:

Pembersihan Data: Menghapus entri dengan nilai hilang.

Normalisasi: Fitur numerik dinormalisasi menggunakan Min-Max Scaling agar nilai berada pada rentang 0 hingga 1.

Binarisasi Kelas: Data label dikonversi menjadi format biner dengan menggabungkan kelas 1–4 menjadi 1 (positif).

#### Prosedur Penelitian

Penelitian ini dilakukan melalui langkah-langkah berikut:

Pseudocode Proses Penelitian:

plaintext

Copy

Edit

1. Import dataset dari sumber UCI
2. Bersihkan data dan lakukan normalisasi
3. Bagi dataset menjadi data latih (80%) dan data uji (20%)
4. Latih model menggunakan algoritma;
  - a. Decision Tree
  - b. Random Forest
  - c. K-Nearest Neighbors
5. Uji model menggunakan data uji
6. Evaluasi performa model berdasarkan metrik:
  - a. Akurasi
  - b. Precision
  - c. Recall
  - d. F1-Score
  - e. Confusion Matrix
7. Bandingkan hasil dan simpulkan algoritma terbaik
  - a. Pengujian dan Evaluasi Model  
Pengujian dilakukan menggunakan data uji yang telah dipisahkan dari data latih menggunakan metode Hold-Out dengan rasio 80:20. Setiap model diuji berdasarkan:  
Accuracy: Proporsi prediksi benar dari total data uji.  
Precision: Kemampuan model untuk menghindari false positive.  
Recall: Kemampuan model untuk mendeteksi seluruh kasus positif (penting dalam konteks medis).  
F1-Score: Harmonik rata-rata dari precision dan recall.  
Confusion Matrix: Visualisasi jumlah prediksi benar dan salah untuk masing-masing kelas.  
Evaluasi dilakukan menggunakan fungsi `classification_report()` dan `confusion_matrix()` dari pustaka `scikit-learn`.

#### Referensi

- [3] Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., et al. (1988). Cleveland Heart Disease Dataset. UCI Machine Learning Repository.  
<https://archive.ics.uci.edu/ml/datasets/heart+disease>

#### 3. Hasil dan Pembahasan

Penelitian ini membandingkan tiga metode klasifikasi dalam mendiagnosis penyakit jantung: Decision Tree, Random Forest, dan K-Nearest Neighbors (KNN). Ketiga algoritma diuji menggunakan data uji sebanyak 60 entri dan dievaluasi menggunakan metrik akurasi, precision, recall, dan f1-score. Hasil pengujian ditunjukkan pada Tabel 1.

Tabel 1. Hasil Evaluasi Kinerja Tiga Model Klasifikasi

Penulis Pertama II Tiga Kata Pertama dari Judul Paper ...

```

Decision Tree
Accuracy: 0.6833333333333333
Confusion Matrix:
[[24  8]
 [31 17]]
Classification Report:
      precision    recall  f1-score   support

 0       0.69      0.75      0.72         32
 1       0.68      0.61      0.64         28

 accuracy      0.68      0.68      0.68         60
 macro avg     0.68      0.68      0.68         60
 weighted avg   0.68      0.68      0.68         60

Random Forest
Accuracy: 0.8666666666666667
Confusion Matrix:
[[24  8]
 [ 5 23]]
Classification Report:
      precision    recall  f1-score   support

 0       0.85      0.91      0.88         32
 1       0.88      0.82      0.85         28

 accuracy      0.87      0.86      0.87         60
 macro avg     0.87      0.86      0.87         60
 weighted avg   0.87      0.87      0.87         60

K-Nearest Neighbors
Accuracy: 0.8833333333333333
Confusion Matrix:
[[26  2]
 [ 5 23]]
Classification Report:
      precision    recall  f1-score   support

 0       0.86      0.94      0.90         32
 1       0.92      0.82      0.87         28

 accuracy      0.89      0.88      0.88         60
 macro avg     0.89      0.88      0.88         60
 weighted avg   0.89      0.88      0.88         60

```

Model	Accuracy	Precision (0/1)	Recall (0/1)	F1-Score (0/1)
Decision Tree	0.683	0.69 / 0.68	0.75 / 0.61	0.72 / 0.64
Random Forest	0.867	0.85 / 0.88	0.91 / 0.82	0.88 / 0.85
K-Nearest Neighbor		0.883	0.86 / 0.92	0.94 / 0.82
			0.87	0.90 / 0.87

Berdasarkan Tabel 1, model K-Nearest Neighbors (KNN) menunjukkan performa terbaik dengan akurasi tertinggi sebesar 88.33%, diikuti oleh Random Forest dengan akurasi 86.67%, dan Decision Tree dengan akurasi 68.33%.

Model Decision Tree memiliki kelemahan dalam mengenali pasien yang sebenarnya sakit (kelas 1), ditunjukkan oleh nilai recall sebesar 0.61. Artinya, sebanyak 39% pasien yang sakit tidak dikenali dengan benar oleh model. Nilai f1-score untuk kelas 1 juga relatif rendah yaitu 0.64, menandakan ketidakseimbangan antara precision dan recall. Meskipun Decision Tree mudah diinterpretasi, hasilnya tidak cukup andal untuk tugas deteksi medis yang kritis.

Sementara itu, Random Forest sebagai metode ensemble yang terdiri dari banyak pohon keputusan menghasilkan performa lebih stabil dan akurat. Recall untuk kelas 1 sebesar 0.82 menunjukkan kemampuan yang baik dalam mendeteksi pasien yang sakit, meskipun masih memiliki 5 false negative.

Model KNN secara keseluruhan menghasilkan nilai f1-score tertinggi baik pada kelas 0 (sehat) maupun kelas 1 (sakit), masing-masing sebesar 0.90 dan 0.87. Precision kelas 1 mencapai 0.92, menandakan bahwa hampir semua prediksi pasien sakit memang benar. Meskipun recall pada kelas 1 sama dengan Random Forest (0.82), nilai precision yang lebih tinggi menjadikan KNN sebagai model paling efektif dalam kasus ini.

Berdasarkan Gambar 1, KNN dan Random Forest keduanya menghasilkan jumlah false negative sebanyak 5, tetapi KNN hanya menghasilkan 2 false positive, lebih sedikit dibandingkan Random Forest yang menghasilkan 3. Decision Tree menghasilkan 8 false positive dan 11 false negative, yang memperkuat temuan bahwa performanya kurang optimal.

#### 4. Kesimpulan

Berdasarkan hasil evaluasi terhadap tiga algoritma klasifikasi yaitu Decision Tree, Random Forest, dan K-Nearest Neighbors (KNN), diperoleh bahwa algoritma K-Nearest Neighbors menunjukkan performa terbaik dengan akurasi sebesar 88.33%, diikuti oleh Random Forest dengan 86.67%, dan Decision Tree dengan 68.33%. Hal ini menunjukkan bahwa metode KNN lebih andal dalam mengklasifikasikan data pada studi ini, sebagaimana tercermin dari nilai precision, recall, dan f1-score yang secara konsisten lebih tinggi dibandingkan metode lainnya.

Temuan ini selaras dengan harapan yang telah dikemukakan pada bab Pendahuluan dan Hasil dan Diskusi, di mana pemilihan metode klasifikasi yang efektif dan akurat sangat penting untuk meningkatkan kualitas pengambilan keputusan berbasis data. KNN terbukti memberikan hasil yang paling mendekati ideal dengan ketidakseimbangan yang minim dalam Confusion Matrix dan skor evaluasi yang stabil.

Prospek pengembangan ke depan, hasil penelitian ini dapat diperluas dengan:

Menggunakan teknik normalisasi atau scaling untuk meningkatkan performa model lebih lanjut, terutama pada KNN.

Menerapkan metode optimasi parameter (hyperparameter tuning) untuk Random Forest dan KNN guna memperoleh kinerja yang lebih maksimal.

Mencoba algoritma klasifikasi lain seperti SVM atau Gradient Boosting sebagai pembanding.

Mengimplementasikan model ke dalam sistem nyata untuk klasifikasi secara otomatis dan evaluasi terhadap data real-time.

Dengan demikian, penelitian ini tidak hanya memenuhi ekspektasi awal, tetapi juga memberikan landasan kuat bagi pengembangan dan penerapan model klasifikasi dalam skenario yang lebih luas di masa mendatang.

#### Referensi

- [1] V. Kumar, "Healthcare Analytics Made Simple". Packt Publishing, 2018.
- [2] Eduonix Learning Solutions, "Machine Learning for Healthcare Analytics Projects". Packt Publishing, 2018.
- [3] DeTrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., et al. (1988). Cleveland Heart Disease Dataset. UCI Machine Learning Repository.

Commented [U1]: Kesimpulan merupakan jawaban tujuan penelitian dan bukan rangkuman dari hasil-hasil yang diperoleh

ORIGINALITY REPORT

---

30%	24%	17%	15%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

---

PRIMARY SOURCES

---

1	Submitted to kptsi2023-1 Student Paper	6%
2	ejournal-binainsani.ac.id Internet Source	2%
3	publikasi.mercubuana.ac.id Internet Source	2%
4	Indra irawan, Wardianto Wardianto, M.Hizbul Wathan, M. Bagus Prayogi. "Studi Perbandingan: Algoritma Random Forest, Naive Bayes Dan Support Vector Machine Dalam Analisis Sentimen Pada Aplikasi Capcut Di Google Play Store", Jurnal Pengembangan Sistem Informasi dan Informatika, 2024 Publication	1%
5	Submitted to Universitas Bina Sarana Informatika Student Paper	1%
6	academic.oup.com Internet Source	1%
7	media.neliti.com Internet Source	1%
8	www.mdpi.com Internet Source	1%

---

repository.unsri.ac.id

9	Internet Source	1 %
10	Submitted to Universitas Airlangga Student Paper	1 %
11	Submitted to Universitas Islam Bandung Student Paper	1 %
12	Submitted to University of Essex Student Paper	1 %
13	<a href="http://www.tatestreetart.com">www.tatestreetart.com</a> Internet Source	1 %
14	Submitted to Universitas Pamulang Student Paper	1 %
15	Yusrinnatul Jinana Triadin, Kusrini Kusrini, Kusnawi Kusnawi. "KLASIFIKASI RANDOM FOREST TERHADAP DIAGNOSA PENYAKIT KANKER PAYUDARA BERDASARKAN STATUS KEGANASAN", TEKNIMEDIA: Teknologi Informasi dan Multimedia, 2025 Publication	1 %
16	<a href="http://repository.nusamandiri.ac.id">repository.nusamandiri.ac.id</a> Internet Source	1 %
17	A. Reza Baehaqa Jamroni Jamroni, Wahyu Hadikristanto, Muhamad Fatchan. "Analisis Faktor dan Prediksi Atrisi untuk Optimalisasi Retensi Karyawan Menggunakan Machine Learning", bit-Tech, 2025 Publication	1 %
18	<a href="http://arxiv.org">arxiv.org</a> Internet Source	1 %



19	Internet Source	1 %
20	<a href="http://www.ijraset.com">www.ijraset.com</a> Internet Source	1 %
21	Submitted to Universitas Pendidikan Ganesha Student Paper	1 %
22	<a href="http://ejournal.akakom.ac.id">ejournal.akakom.ac.id</a> Internet Source	1 %
23	Ivan Aditya Nugraha, Irving Vitra Paputungan. "Analisis Sentimen Video Review Starlink Indonesia Menggunakan Pendekatan Lexicon Dictionary Bahasa Indonesia Dan Algoritma Random Forest", Innovative: Journal Of Social Science Research, 2025 Publication	<1 %
24	Kang Andini Wulandari, Adhitya Nugraha, Ardytha Luthfiarta, Laila Rahmatin Nisa. "Peningkatan Akurasi Deteksi Dini Penyakit Parkinson melalui Pendekatan Ensemble Learning dan Seleksi Fitur Optimal", Edumatic: Jurnal Pendidikan Informatika, 2024 Publication	<1 %
25	Marwa Sulehu, Wisda Wisda, First Wanita, Markani Markani. "Optimasi Prediksi Kelulusan Mahasiswa Menggunakan Random Forest untuk Meningkatkan Tingkat Retensi", Jurnal Minfo Polgan, 2025 Publication	<1 %
26	Yonal Supit. "Penerapan Algoritma Support Vector Machine (SVM) untuk Deteksi Penyakit	<1 %

# Mata Katarak", INDONESIAN JOURNAL ON DATA SCIENCE, 2025

Publication

27

[ojs.uho.ac.id](https://ojs.uho.ac.id)

Internet Source

<1 %

28

[pt.scribd.com](https://pt.scribd.com)

Internet Source

<1 %

29

[text-id.123dok.com](https://text-id.123dok.com)

Internet Source

<1 %

30

Fitri Handayani. "Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network Dalam Prediksi Penyakit Jantung", Jurnal Edukasi dan Penelitian Informatika (JEPIN), 2021

Publication

<1 %

31

Nasrah, Korinus Suweni, Sofitje J Gentindatu, Kismiyati, Marjuannah. "PENGEMBANGAN MODEL SELF-ESTEEM UNTUK MENINGKATKAN KUALITAS HIDUP PASIEN TUBERKULOSIS PARU: STUDI KUALITATIF", JURNAL KEPERAWATAN TROPIS PAPUA, 2024

Publication

<1 %

Exclude quotes

Off

Exclude matches

Off

Exclude bibliography

Off