



Implementasi Algoritma Decision Tree untuk Klasifikasi Serangan Jantung

*¹ Muhammad Akram Fais, ² M. Revano Ananda Lubis, ³ Annisa Aulia, ⁴ Indri Syafitri

^{1,2,3} Ilmu Komputer, Fakultas Matematika, ⁴ Ilmu Pengetahuan Alam, Universitas Negeri Medan

E-mail: ¹ mhdakramfais@mhs.unimed.ac.id, ² revanoananda@mhs.unimed.ac.id, ³

annisaauliaaa@mhs.unimed.ac.id, ⁴ indrisyafitri@mhs.unimed.ac.id

* Korespondensi penulis: mhdakramfais@mhs.unimed.ac.id

Abstract. As many as 7.3 million people worldwide die from heart disease. This indicates that heart disease is one of the diseases that cause the most deaths. As a preventive effort in handling heart disease, it is necessary to predict heart disease in patients. The classification process to predict heart disease is done using a decision tree. This decision tree is interesting because it is more flexible in providing the advantage of visualizing the advice so that the prediction can be observed. This study uses Heart Disease Prediction Dataset data with a total of 303 data. Then predictions are made using Decision tree so that the accuracy results are 83.60%, precision 89.28%, recall 78.12% and F1 score of 83.33%.

Keywords: heart disease, Prediction, Decision Tree

Abstrak. Sebanyak 7,3 juta orang di seluruh dunia meninggal disebabkan karena serangan jantung. Hal ini menandakan bahwa serangan jantung merupakan salah satu hal yang paling banyak menyebabkan kematian. Untuk mencegah serangan jantung, penting untuk melakukan prediksi potensi serangan jantung pada pasien sebagai tindakan pencegahan. Proses klasifikasi menggunakan *Decision Tree* digunakan untuk memprediksi kemungkinan serangan jantung. *Decision tree* ini menarik karena memberikan fleksibilitas yang lebih tinggi dalam memberikan manfaat berupa representasi visual saran, sehingga prediksinya dapat diamati. Data yang digunakan dalam penelitian ini adalah data *Heart Disease Prediction Dataset* dengan jumlah data sebanyak 303. Kemudian dilakukan prediksi menggunakan *Decision tree* sehingga didapat hasil akurasi sebesar 83.60%, *precision* 89.28%, *recall* 78.12% dan *F1 score* sebesar 83.33%.

Kata kunci: Serangan Jantung, Prediksi, *Decision Tree*

PENDAHULUAN

Salah satu organ utama dalam peredaran darah pada tubuh manusia adalah jantung. Fungsi jantung dapat terhambat apabila tidak menjaga kesehatan seperti, merokok, minim beraktivitas, stres berat, kolesterol yang tinggi, dan juga obesitas. Menurut informasi yang dikumpulkan oleh WHO, serangan jantung menjadi penyebab fatal bagi 7,3 juta nyawa di berbagai belahan dunia. Karena jantung tidak dapat melakukan fungsinya dengan baik, seperti otot jantung yang lemah, tipe serangan jantung muncul [1]. Serangan jantung terjadi ketika jantung bekerja dengan buruk, menghambat pemompaan darah dan aliran oksigen ke seluruh tubuh [2]. Serangan jantung adalah kondisi yang sangat serius, dan bila tidak ditangani dengan cepat, dapat mengakibatkan kematian. Ini adalah suatu gangguan jantung di mana otot jantung mengalami kekurangan aliran darah, sehingga menghambat kemampuan jantung untuk memompa darah ke seluruh tubuh. Serangan jantung sering kali sulit untuk dideteksi secara dini, sehingga memerlukan pendekatan yang efektif dalam pengklasifikasian serangan jantung sehingga petugas medis dapat dengan lebih efisien menemukan tanda-tanda serangan jantung

pada pasien. Satu contohnya adalah penerapan teknologi pembelajaran mesin. Klasifikasi merupakan proses pengelompokan item yang memiliki ciri-ciri serupa ke dalam beberapa kelas[3]. Dalam kebanyakan kasus, pengklasifikasian diwakili dengan kalimat-kalimat penting yang menunjukkan karakteristik atau atributnya. Terdapat banyak model dan konsep dalam *data mining*, di antaranya adalah *Decision Tree*. Pohon tersebut diterapkan dalam kecerdasan untuk mencari solusi masalah, dan pohon tersebut tidak terbatas pada struktur biner saja. Metode ini menarik karena memberikan fleksibilitas tinggi dalam menyajikan keuntungan melalui visualisasi saran sehingga prediksinya dapat diamati. Namun, penggunaan sejumlah besar kelas dan kriteria dapat menyebabkan keputusan menjadi lebih lambat dan memerlukan kapasitas memori yang besar [4].

Penelitian terdahulu dengan menggunakan *Decision Tree* seperti, pengaplikasian *Decision Tree Biner* pada penyakit gigi[5], Klasifikasi Kualitas Berita Pada Majalah Menggunakan Metode *Decision Tree* [6], pengeplikasian *Decision Tree* dalam klasifikasi gizi anak[7], Implementasi *Decision Tree* untuk Klasifikasi Kompetensi Siswa [8]. Berdasarkan penelitian di atas, penelitian ini menggunakan algoritma decision tree untuk menentukan hasil klasifikasi serangan jantung. Dengan menggunakan algoritma decision tree penelitian ini bertujuan untuk mendapatkan hasil yang akurat, sehingga dalam bidang medis dapat menggunakan informasi klasifikasi ini ntuk mengidentifikasi pasien yang berisiko tinggi dan mengambil tidakan tepat.

METODE PENELITIAN

1 Sumber Data

Dalam penelitian ini, sumber data yang digunakan berasal dari data sekunder yang diunduh dari situs web Kaggle dengan judul "*Heart Disease Prediction*," yang terdiri dari 303 entri [9]. Terdapat 13 atribut yang diperlihatkan pada tabel 1.

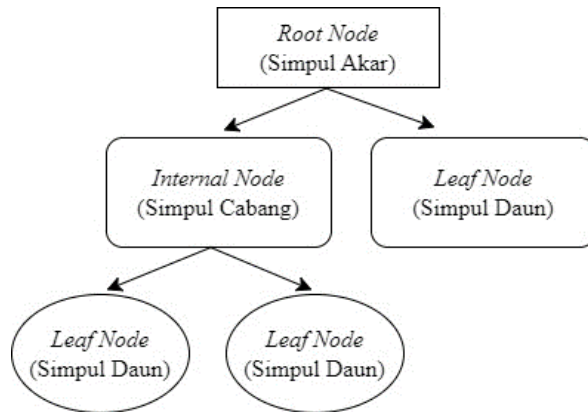
Tabel 1. Daftar Atribut Dataset

Atribut	Keterangan
Age	Usia
Sex	Jenis Kelamin
Cp	Rasa sakit pada area dada
Trestbps	Tekanan darah saat tidak beraktivitas
Chol	Kolesterol
Fbs	Gula darah saat berpuasa
restecg	Hasil elektrokardiografi saat istirahat.
thalach	Detak jantung maksimal
exang	Latihan yang diinduksi angina

Oldpeak	Depresi yang diinduksi oleh latihan <i>relative</i>
Slope	Kemiringan puncak ST Segmen
caa	Jumlah pembuluh darah yang berwarna setelah diwarnai flourosopy
Thall	Tipe kerusakan pembuluh darah
output	Indikasi serangan jantung

2 *Decision Tree*

Decision Tree merupakan model prediktif dalam *machine learning* yang digunakan untuk membuat keputusan berdasarkan berbagai aturan. Model ini mirip dengan struktur pohon keputusan, dimana setiap simpul pada pohon mewakili hasil pengujian dan setiap daun mewakili nilai prediksi. *Decision Tree* dapat digunakan untuk klasifikasi dan pemodelan regresi. Struktur *Decision Tree* dapat dilihat pada gambar 1.



Gambar 1. Struktur *Decision Tree*

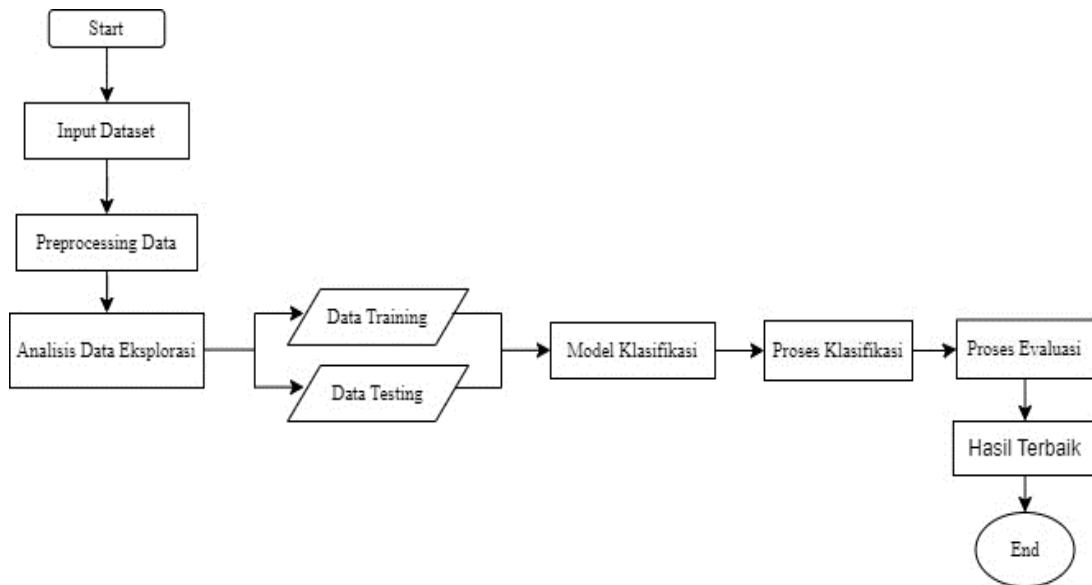
Penilaian tingkat efektivitas suatu atribut dalam proses klasifikasi disebut sebagai *information gain*, yang dihitung berdasarkan nilai *gain* tertinggi sesuai dengan persamaan 1[10].

$$Gain(S, A) = Entropy(S) - \sum v \in Values(A) \frac{|S_v|}{|S|} Entropy(S_v).....(1)$$

dimana *S* adalah himpunan kasus, *A* adalah atribut, *N* adalah sejumlah partisi atribut *A*, *|S_v|* merupakan jumlah kasus dalam partisi ke-*i*, dan *|S|* adalah jumlah kasus dalam *S*. Perhitungan nilai Entropy dilakukan sebelum mendapatkan nilai *gain*. Persamaan untuk *Entropy* dijelaskan pada persamaan 2.

$$Entropy(S) = -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus}.....(2)$$

3 Langkah Penelitian



Gambar 2. Alur Penelitian

Dalam gambar 2, tergambar urutan langkah-langkah penelitian yang melibatkan beberapa tahapan, seperti pengisian nilai yang hilang dalam dataset, melaksanakan preprocessing data, menjalankan klasifikasi menggunakan *Decision Tree*, dan mengukur nilai evaluasi guna memperoleh hasil klasifikasi.

HASIL DAN PEMBAHASAN

Pada bagian ini disajikan sesuai penelitian yang dilakukan. Pengujian yang dilakukan menggunakan algoritma *Decision Tree* menghasilkan klasifikasi pada Serangan Jantung. Dalam pengujian ini menggunakan 13 feature untuk mengklasifikasi Serangan Jantung.

1 Data Preprocessing

Sebelum data dimodelkan ke dalam algoritma *Decision Tree*, data perlu dilakukan *cleansing* dan *preprocessing* terlebih dahulu agar data yang dimodelkan bersih dari data yang hilang dan memiliki data yang bersih.

Tabel 2. Pengecekan Data Hilang

Atribut	Data Hilang
Age	Tidak ada
Sex	Tidak ada
Cp	Tidak ada
Trtbps	Tidak ada
Chol	Tidak ada
Fbs	Tidak ada

restecg	Tidak ada
Thalachh	Tidak ada
Exng	Tidak ada
Oldpeak	Tidak ada
Slp	Tidak ada
Caa	Tidak ada
Thall	Tidak ada
output	Tidak ada

Berdasarkan tabel 2, tidak ada nilai yang hilang pada dataset sehingga tidak perlu dilakukan penghapusan ataupun penambahan data.

3.2 Implementasi Decision Tree

Setelah dilakukan preprocessing data, tahap selanjutnya adalah memodelkan data ke dalam algoritma *Decision Tree*. Tabel 3 menunjukkan hasil implementasi pada algoritma *Decision Tree*.

Tabel 3 Performa Decision Tree

Metric	Tidak Rentan	Rentan	AVG	AVG PERF
Acc	83%	83%	83%	83,63%
Prec	79%	89%	84%	
Rec	90%	78%	84%	
F1 Score	84%	83%	83,5%	

Berdasarkan analisis kinerja yang dilakukan algoritma *Decision Tree* dengan splitting data 80% dan 20%, secara keseluruhan memiliki tingkat performa yang baik yaitu 83,63%. Menurut jenis pengukuran performa, algoritma *Decision Tree* dengan pengukuran performa *Recall* (Sensitifitas) dan *Precission* memiliki tingkat performa yang tertinggi yaitu 84%.

KESIMPULAN

Dalam Penelitian ini, dilakukan implementasi klasifikasi pada Serangan Jantung menggunakan algoritma *Decision Tree*. Berdasarkan hasil yang didapat dari pengolahan data, penggunaan algoritma *Decision Tree* dalam klasifikasi Serangan Jantung menghasilkan akurasi sebesar 83,63%. Dengan akurasi 83,63% algoritma *Decision Tree* sudah dapat dikatakan akurat dalam mengklasifikasi Serangan Jantung.

SARAN

Hasil akurasi dari algoritma *Decision Tree* untuk klasifikasi serangan jantung mencapai 83,63% hal ini menunjukkan bahwa algoritma *Decision Tree* sudah dapat dikatakan akurat.

Akan tetapi akurasi masih memungkinkan untuk ditingkatkan lagi, salah satunya dengan cara memotong fitur yang kurang berpengaruh dalam proses pemodelan algoritma *Decision Tree*. Algoritma klasifikasi lainnya juga dapat digunakan dan dibandingkan dengan algoritma *Decision Tree* yang memiliki kemungkinan akurasi lebih tinggi.

DAFTAR PUSTAKA

- [1] D. Haganta Depari *et al.*, “Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Serangan Jantung,” *JURNAL INFORMATIK Edisi ke*, vol. 18, p. 2022.
- [2] J. Khatib Sulaiman, A. A. Mizwar Rahim, I. Yanuar Risca Pratiwi, M. Ainul Fikri, and U. Amikom Yogyakarta, “Klasifikasi Serangan Jantung Menggunakan Metode Synthetic Minority Over-Sampling Technique Dan Random Forest Classifier,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 5, pp. 2023–2995.
- [3] J. Homepage *et al.*, “MALCOM: Indonesian Journal of Machine Learning and Computer Science Implementation of Decision Tree Algorithm and Support Vector Machine for Lung Cancer Classification Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Serangan Kanker Paru,” vol. 3, pp. 15–19, 2023.
- [4] P. Kurnia Illahi, A. Rina Viana, N. Fitria, M. Permata, and M. Y. Pratama, “SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Application of Decision Tree Algorithm and Linear Regression for Breast Cancer Classification Penerapan Algoritma Decision Tree aan Regresi Linear untuk Klasifikasi Kanker Payudara.” [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [5] M. Ula, F. T. T. Anjani, A. F. Ulva, I. Sahputra, and A. Pratama, “APPLICATION OF MACHINE LEARNING WITH THE BINARY DECISION TREE MODEL IN DETERMINING THE CLASSIFICATION OF DENTAL DISEASE,” *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 6, no. 1, pp. 170–179, Jul. 2022, doi: 10.31289/jite.v6i1.7341.
- [6] A. Irma Purnamasari and A. Rinaldi Dikananda, “Klasifikasi Kualitas Berita Pada Majalah Menggunakan Metode Decision Tree,” *Jurnal Teknologi Ilmu Komputer*, vol. 1, no. 2, pp. 48–54, 2023, doi: 10.56854/jtik.v1i2.52.
- [7] M. Ula, A. F. Ulva, M. Mauliza, M. A. Ali, and Y. R. Said, “APPLICATION OF MACHINE LEARNING IN DETERMINING THE CLASSIFICATION OF CHILDREN’S NUTRITION WITH DECISION TREE,” *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 5, pp. 1457–1465, Sep. 2022, doi: 10.20884/1.jutif.2022.3.5.599.
- [8] H. Rifa, R. Hamonangan, and D. Ade Kurnia, “KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer Implementasi Algoritma Decision Tree Dalam Klasifikasi Kompetensi Siswa”, [Online]. Available: <http://jurnal.kopertipindonesia.or.id/>
- [9] Bilal Husain, M “Heart Disease Prediction” September. 2023. Available: <https://www.kaggle.com/code/bilalchanna/heart-disease-prediction> (accessed Nov 20, 2023).
- [10] Mitchell, T. M. 1997. *Machine Learning*. McGraw-Hill.