

KLASIFIKASI RISIKO PENYAKIT SERANGAN JANTUNG DENGAN MENGUNAKAN ALGORITMA C4.5

Aida Khalisatifa¹⁾, Hestiana Dela Arum²⁾, Muhammad Ihsan Jambak³⁾

^{1,2,3)}Universitas Sriwijaya

Email : akhalisatf@gmail.com ¹⁾, hestianadaa@gmail.com ²⁾, jambak@unsri.ac.id ³⁾

ABSTRAK

Penyakit jantung termasuk serangan jantung merupakan salah satu penyakit yang menjadi perhatian serius dalam dunia medis. Deteksi dini risiko penyakit jantung sangat penting untuk mencegah dampak yang lebih serius. Penelitian ini menggunakan dataset yang terdiri dari 303 data dengan tipe numerik dan nominal. Data tersebut diolah menggunakan algoritma C4.5 untuk melakukan klasifikasi risiko penyakit serangan jantung. Hasil penelitian menunjukkan bahwa algoritma tersebut mencapai tingkat akurasi sebesar 83,98%. Lebih lanjut, melalui decision tree mengidentifikasi bahwa faktor terbesar yang diketahui berkontribusi dalam penentuan risiko penyakit serangan jantung adalah faktor Cp (Chest Pain), Caa (n Major Vessels), Oldpeak, Sex, dan Exng (Exercise Angina). Temuan penelitian ini diharapkan dapat memberikan tambahan wawasan dalam upaya deteksi dini dan penanganan penyakit jantung di masa mendatang di mana model klasifikasi ini dapat digunakan sebagai alat skrining awal untuk mengidentifikasi individu yang berisiko tinggi terkena serangan jantung.

Kata Kunci : Klasifikasi, Risiko, Serangan jantung, Algoritma C4.5

ABSTRACT

Heart disease, including heart attacks, is a serious concern in the medical field. Early detection of heart disease risk is crucial to prevent more serious consequences. This study utilized a dataset consisting of 303 records with both numerical and nominal data types. The data were processed using the C4.5 algorithm to classify the risk of heart attack disease. The research findings revealed that the algorithm achieved with accuracy rate of 83,98%. Furthermore, through decision tree analysis, it was identified that the most significant contributing factor to determining the risk of heart attack disease is factor Cp (Chest Pain), Caa (n Major Vessels), Oldpeak, Sex, and Exng (Exercise Angina). The findings of this study are expected to provide additional insights into early detection and management of heart disease in the future, where this classification model can serve as an early screening tool to identify individuals at high risk of heart attack.

Keywords: Classification, Risk, Heart attack, C4.5 Algorithm

1. PENDAHULUAN

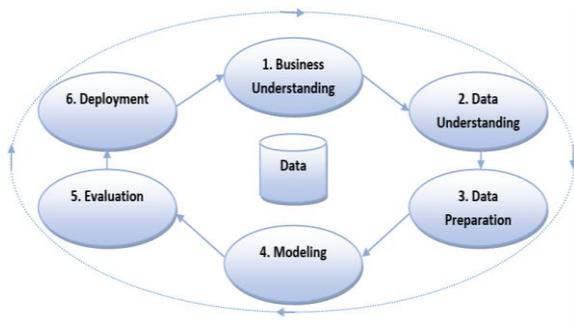
Serangan jantung merupakan kondisi dimana aliran darah yang membawa oksigen ke otot jantung tiba-tiba tersumbat. Sehingga, jantung tidak akan mendapat cukup oksigen yang dapat berakibat pada rusaknya otot jantung. Penyebab umum penyumbatan mendadak dikarenakan pembentukan bekuan darah (thrombus). Bekuan darah terbentuk akibat adanya atherosclerosis, suatu kondisi di mana timbunan lemak (plak) menumpuk di dinding bagian dalam pembuluh darah (*Heart attack (myocardial infarction)*, 2023). Sebagian besar serangan jantung disebabkan oleh atherosclerosis. Namun, faktor usia, gaya hidup yang tidak sehat, serta kondisi medis lainnya dapat meningkatkan resiko terkena serangan jantung (*What Is a Heart Attack*, 2022). Penyakit jantung, termasuk serangan jantung, adalah salah satu penyebab kematian utama di seluruh dunia. Berdasarkan data per 2021 yang dikeluarkan oleh WHO (World Health Organization), organisasi internasional yang berfokus pada kesehatan global, menyatakan bahwa penyakit jantung menjadi penyebab yang merenggut 17,9 juta nyawa di dunia. Terutama, serangan jantung dan stroke yang berkontribusi lebih dari 80% kematian akibat penyakit jantung (*Cardiovascular diseases*).

Diketahui bahwa penyakit jantung menjadi permasalahan yang cukup serius. Seseorang belum tentu waspada dan mengetahui apakah dirinya berpotensi untuk mengidap penyakit tersebut. Untuk itu, deteksi secara dini diperlukan setiap individu agar resiko tersebut dapat dicegah ataupun diobati. Dengan perkembangan teknologi yang semakin canggih, data mining dapat digunakan sebagai salah satu alternatif untuk membantu prediksi penyakit jantung. Dengan menggunakan algoritma C4.5 bisa membantu klasifikasi untuk menentukan apakah seseorang tersebut memiliki resiko serangan jantung atau tidak. Metode Decision Tree C4.5 digunakan untuk membangun model prediksi kelas objek dengan memilih variabel yang signifikan, yang sering digunakan dalam penelitian di bidang kesehatan dan kedokteran (Arkamil and Jambak, 2023).

Berdasarkan penelitian yang berjudul “Klasifikasi Jantung Dengan Menggunakan Algoritma C4.5” yang bertujuan untuk memprediksi penyakit jantung dengan menggunakan algoritma C4.5, hasil yang diperoleh menunjukkan tingkat akurasi sebesar 79% (Sepharni, Hendrawan and Rozikin, 2022). Selanjutnya, penelitian berjudul “Klasifikasi Penyakit Gagal Jantung Menggunakan Supervised Learning” bertujuan untuk membandingkan performa klasifikasi penyakit gagal jantung dengan menggunakan tiga algoritma berbeda, untuk menentukan algoritma yang memiliki tingkat akurasi paling optimal dalam pengujian data penyakit gagal jantung. Penelitian tersebut mengungkapkan bahwa algoritma C4.5 mencapai nilai akurasi tertinggi, yakni 80% (Tasia *et al.*, 2023). Lalu, penelitian yang berjudul “Klasifikasi Penyakit Jantung Menggunakan Algoritma Decision Tree Series C4.5 Dengan Rapidminer” dalam melakukan proses klasifikasi data penyakit jantung berhasil memperoleh tingkat akurasi sebesar 80,34% (Agus Oka Gunawan *et al.*, 2023). Penelitian-penelitian tersebut memiliki kesamaan dengan algoritma yang digunakan dan fokus bidang penelitian yaitu untuk penyakit jantung. Maka, berdasarkan hal tersebut, penelitian ini akan melihat keefektifan algoritma C4.5 dalam melakukan klasifikasi serangan jantung menggunakan dataset yang berbeda.

2. METODE

Alur pada penelitian ini, mengadopsi standardisasi dalam *data mining*, yaitu CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang terbagi menjadi 6 tahap, meliputi *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Adapun, tahapan dalam CRISP-DM diilustrasikan seperti pada Gambar 1 berikut:



Gambar 1. Tahapan Model CRISP-DM (North, 2016)

2.1. Business Understanding

Pada tahap *business understanding*, dapat dilakukan analisis masalah untuk memahami permasalahan yang telah ditemukan. Penelitian ini ditujukan untuk mengetahui faktor-faktor atau karakteristik yang menjadi penyebab dari risiko penyakit serangan jantung, sehingga setiap individu dapat lebih waspada dan *aware* akan potensi untuk mengidap penyakit tersebut.

2.2. Data Understanding

Pada penelitian ini menggunakan dataset *Heart Attack Analysis* yang bersifat publik, diakses melalui situs Kaggle, dan terdiri dari 14 atribut dengan 303 *record* data. Atribut Output berperan sebagai label pada dataset ini. Adapun, informasi mengenai dataset seperti nama atribut, deskripsi, dan tipe data dapat dilihat pada Tabel 1 di bawah ini:

Tabel 1. Informasi Dataset

Atribut	Deskripsi	Tipe Data
age	Umur dalam tahun	Numerik
sex	Jenis kelamin (0 = wanita, 1 = pria)	Biner
cp	Jenis nyeri dada yang dirasakan, dengan nilai kategorikal untuk jenis tertentu	Nominal
trtbps	Tekanan darah saat istirahat, diukur dalam mmHg	Numerik

chol	Kadar kolesterol serum, diukur dalam mg/dL	Numerik
fbbs	Kategori apakah gula darah puasa lebih dari 120 mg/dL	Biner
restecg	Hasil elektrokardiografi saat istirahat, dengan kategori khusus untuk berbagai hasil	Nominal
thalachh	Detak jantung maksimum yang tercapai selama tes stres, diukur dalam BPM	Numerik
exng	Kategori apakah angina dipicu oleh latihan	Biner
oldpeak	Tingkat depresi segmen ST yang diinduksi oleh latihan relatif terhadap istirahat	Numerik
slp	Kemiringan segmen ST puncak selama tes latihan, dengan kategori untuk berbagai kemiringan	Nominal
caa	Jumlah pembuluh darah utama yang diberi warna oleh fluoroskopi, dalam rentang 0 hingga 4	Nominal
thall	Hasil tes stres talium, dengan kategori untuk berbagai hasil	Nominal
output	Status risiko penyakit jantung	Biner

2.3. Data Preparation

Pada tahap ini merupakan proses dimana data dipersiapkan agar siap untuk dianalisis dan dimodelkan, yang meliputi

proses menyusun, menggabungkan, serta mengumpulkan data (Damanik and Jambak, 2023). Pada penelitian ini, dilakukan *data transformation* untuk beberapa atribut yang bersifat numerik menjadi kategorikal. Untuk usia (Age) dikategorikan menjadi 5 kategori, sedangkan tekanan darah sistolik (Trtbps), kadar kolesterol (Chol), dan denyut jantung maksimal dikategorikan menjadi 3 kategori (Santoso *et al.*, 2023). Untuk atribut lainnya ditransformasikan sesuai dengan kategori yang telah ditetapkan pada informasi dataset yang terlihat di Tabel 1. Berikut pada Tabel 2 merupakan rincian *record* dataset sebelum dilakukan transformasi data

Tabel 2. Dataset *Heart Attack Analysis* Sebelum *Data Transformation*

age	sex	cp	trtbps	chol	old peak	caa	...	output
63	1	3	145	233	2.3	0	...	1
37	1	2	130	250	3.5	0	...	1
41	0	1	130	204	1.4	0	...	1
56	1	1	120	236	0.8	0	...	1
57	0	0	120	354	0.6	0	...	1
...
57	0	1	130	236	0	1	...	0

Tabel 3 di bawah ini merupakan rincian *record* dataset setelah dilakukan transformasi data. Dalam hal ini, beberapa atribut yang bertipe numerik ditransformasikan menjadi tipe kategorikal.

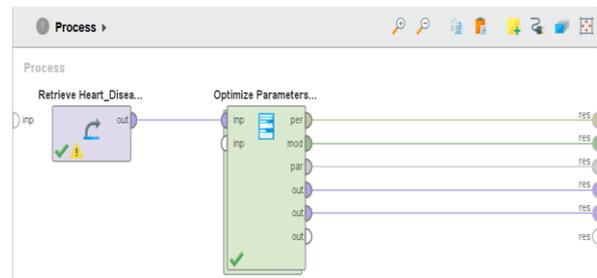
Tabel 3. Dataset *Heart Attack Analysis* Setelah *Data Transformation*

Age	Sex	Cp	Trtbps	Chol	Old peak	Caa	...	Output
Elderly	Male	Asymptomatic	High	Threshold	2.3	0	...	High Risk
Adult	Male	Non-anginal pain	Normal	High	3.5	0	...	High Risk
Adult	Female	Atypical angina	Normal	Threshold	1.4	0	...	High Risk
Pre-elderly	Male	Atypical angina	Normal	Threshold	0.8	0	...	High Risk
Pre-elderly	Female	Typical angina	Normal	High	0.6	0	...	High Risk
...
Pre-elderly	Female	Atypical angina	Normal	Threshold	0	1	...	Low Risk

2.4. Modeling

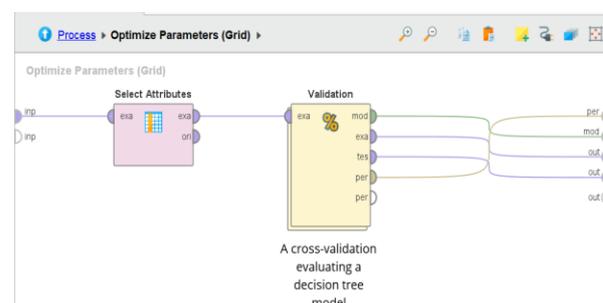
Tahapan berikutnya dalam model proses CRISP-DM adalah *modeling*. Pada tahap ini ditentukan metode *data mining* yang

tepat dan pemilihan algoritma agar sesuai dengan tujuan bisnis dan jenis data. Dalam penelitian ini, peneliti menggunakan metode klasifikasi dengan algoritma *Decision Tree C4.5* yang menerapkan operator *Optimize Parameters (Grid)* seperti pada Gambar 2. Operator ini berfungsi memberikan nilai optimal berdasarkan set parameter yang dipilih untuk Operator dalam sub prosesnya.



Gambar 2. Rancangan Model dengan Operator *Optimize Parameters (Grid)*

Di dalam sub-proses dari Operator *Optimize Parameters (Grid)* diterapkan Operator *Select Attributes* dan Operator *Cross Validation* yang terlihat pada Gambar 3. *Cross Validation* digunakan untuk membagi data ke dalam beberapa *fold* dengan ukuran yang sama menjadi *data training* dan *data testing*, sehingga dapat memastikan kinerja model berjalan dengan baik pada dataset yang berbeda.



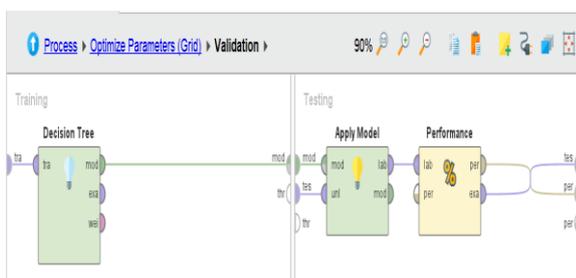
Gambar 3. Operator *Select Attributes* dan *Cross Validation*

Metode klasifikasi diterapkan untuk mengelompokkan data ke dalam kategori atau kelas berdasarkan fitur atau atribut tertentu. Salah satu algoritma klasifikasi yang banyak digunakan adalah C4.5. Algoritma C4.5 merupakan algoritma berbasis *Decision Tree* yang digunakan untuk klasifikasi, segmentasi, atau pengelompokan data dengan membangun

model berbentuk pohon keputusan yang bersifat prediktif (Suherman *et al.*, 2024), dimana setiap *node* pohon keputusan dibuat berdasarkan variabel yang paling berpengaruh dalam memprediksi kategori suatu objek (Arkamil and Jambak, 2023). Adapun, *tools* yang digunakan pada penelitian ini adalah Rapid Miner. Berikut merupakan tahapan secara umum dalam membuat *decision tree* menggunakan algoritma C4.5 (Maulana, Winanjaya and Rizki, 2022):

- Menentukan atribut yang berperan sebagai akar *decision tree*
- Membuat cabang untuk setiap nilai dari atribut
- Membagi *case/data* ke dalam cabang-cabang
- Ulangi tahapan 2 dan 3 untuk setiap cabang hingga semua *case/data* memiliki kelas yang sama.

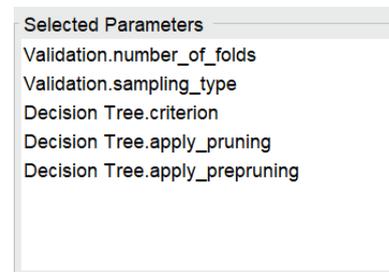
Terdapat sub proses *training* dan *testing* di dalam operator *Cross Validation* seperti yang terlihat pada Gambar 4. Pada sub-proses *training*, dipilih operator *Decision Tree* yang berfungsi untuk melatih dataset agar suatu model dapat dibentuk dan dikembangkan. Kemudian, pada sub-proses *testing* akan dilakukan evaluasi atau pengujian kinerja terhadap model yang telah dibuat. Dengan menerapkan operator *Apply Model* dan *Performance*, maka akurasi dari *testing* dataset juga dapat diketahui.



Gambar 4. Operator *Decision Tree*, *Apply Model*, dan *Performance*

Setelah seluruh operator diterapkan, maka operator *Optimize Parameters (Grid)* akan menampilkan berbagai parameter dari operator *Select Attributes*, *Cross Validation*, *Decision Tree*, *Apply Model*, dan

Performance. Parameter yang dipilih nantinya akan berfokus pada parameter dari operator *Cross Validation* dan *Decision Tree* dengan tujuan untuk menemukan kombinasi set parameter yang memberikan nilai akurasi paling terbaik. Parameter yang dipilih antara lain *number of folds*, *sampling type*, *criterion*, *apply pruning*, dan *apply prepruning* yang dapat dilihat pada Gambar 5.



Gambar 5. Parameter Terpilih pada Operator *Optimize Parameters (Grid)*

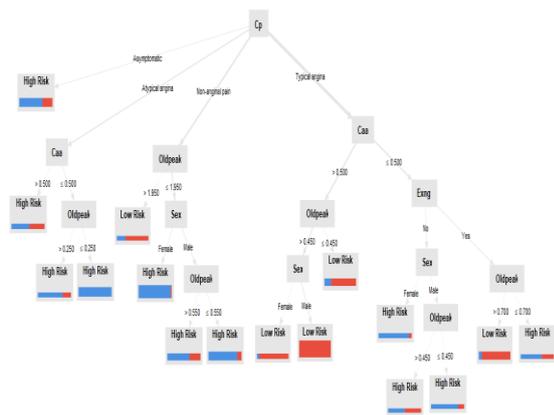
Pada penelitian ini, *Accuracy*, *Precision Recall*, dan *Specificity* menjadi indikator untuk mengevaluasi kinerja model yang dihasilkan. Dengan begitu, klasifikasi dengan algoritma C4.5 dapat menjadi efektif dalam mengidentifikasi individu yang berpotensi berisiko rendah atau tinggi terkena serangan jantung.

3. HASIL DAN PEMBAHASAN

Berdasarkan pengolahan data yang telah dilakukan, diperoleh hasil bahwa proses dari *Optimize Parameters (Grid)* memberikan sebanyak 704 kombinasi parameter dengan nilai akurasi terbaik sebesar 83,98%. Adapun, set parameter yang diperoleh melalui pengolahan *Optimize Parameters (Grid)*, meliputi:

- Number of Folds* = 90
- Sampling Type* = *Automatic*
- Criterion* = *Information gain*
- Apply Pruning* = *False*
- Apply Prepruning* = *True*

Melalui pemodelan yang dilakukan dengan menggunakan algoritma *Decision Tree*, maka diperoleh ilustrasi pola atau *rules* dari dataset *Heart Attack Analysis* yang terlihat pada Gambar 6 berikut:



Gambar 6. Pohon Keputusan *Heart Attack Analysis* Dataset

Berdasarkan ilustrasi pola pohon keputusan yang telah dihasilkan, maka dapat diketahui bahwa terdapat 5 atribut yang menjadi karakteristik utama penyebab dari individu berisiko mengalami penyakit serangan jantung. Atribut tersebut terdiri dari Cp (*Chest Pain*), Caa (*n Major Vessels*), Oldpeak, Sex, dan Exng (*Exercise Angina*). Atribut Cp atau *Chest Pain* berperan sebagai akar (*root*) dari pohon keputusan, dimana faktor *chest pain* perlu diperhatikan pertama kali dalam mengidentifikasi apakah seseorang berisiko mengalami penyakit serangan jantung yang tinggi atau rendah.

Melalui pohon keputusan yang ada, diperoleh aturan atau *rule*, yaitu aturan IF-THEN yang menggambarkan jalur dari akar pohon ke daun. *Rule* menjadi cara yang baik untuk merepresentasikan informasi atau potongan-potongan pengetahuan yang terkandung di dalam pohon keputusan (Han, Kamber and Pei, 2012). Berikut merupakan *rule* yang terbentuk dari pohon keputusan *Heart Attack Analysis* di atas:

Tree

Cp = Asymptomatic: High Risk {High Risk=16, Low Risk=7}

Cp = Atypical angina

| Caa > 0.500: High Risk {High Risk=7, Low Risk=6}

| Caa ≤ 0.500

| | Oldpeak > 0.250: High Risk {High Risk=9, Low Risk=3}

| | Oldpeak ≤ 0.250: High Risk {High Risk=25, Low Risk=0}

Cp = Non-anginal pain

| Oldpeak > 1.950: Low Risk {High Risk=3, Low Risk=8}

| Oldpeak ≤ 1.950

| | Sex = Female: High Risk {High Risk=34, Low Risk=1}

| | Sex = Male

| | | Oldpeak > 0.550: High Risk {High Risk=12, Low Risk=6}

| | | Oldpeak ≤ 0.550: High Risk {High Risk=20, Low Risk=3}

Cp = Typical angina

| Caa > 0.500

| | Oldpeak > 0.450

| | | Sex = Female: Low Risk {High Risk=1, Low Risk=12}

| | | Sex = Male: Low Risk {High Risk=0, Low Risk=46}

| | Oldpeak ≤ 0.450: Low Risk {High Risk=4, Low Risk=15}

| Caa ≤ 0.500

| | Exng = No

| | | Sex = Female: High Risk {High Risk=11, Low Risk=1}

| | | Sex = Male

| | | | Oldpeak > 0.450: High Risk {High Risk=5, Low Risk=5}

| | | | Oldpeak ≤ 0.450: High Risk {High Risk=9, Low Risk=2}

| | Exng = Yes

| | | Oldpeak > 0.700: Low Risk {High Risk=2, Low Risk=19}

| | | Oldpeak ≤ 0.700: High Risk {High Risk=7, Low Risk=4}

Berdasarkan *rule* yang di atas, diperoleh pemahaman bahwa atribut Cp menjadi hal yang pertama ditinjau dari seorang individu. Individu dengan *chest pain* kategori *Asymptomatic*, dikategorikan sebagai *High Risk* dengan jumlah sampel 16 *High Risk* dan

7 *Low Risk*. Individu dengan *chest pain* kategori *Atypical Angina*, terbagi lebih lanjut berdasarkan Caa (n *Major Vessels*) dan Oldpeak. Individu dengan *chest pain* kategori *Non-anginal Pain* dan memiliki Oldpeak lebih dari 1.95, maka dikategorikan sebagai *Low Risk*. Jika Oldpeak lebih rendah dan berjenis kelamin *Female*, maka dikategorikan sebagai *High Risk* dengan jumlah sampel sebanyak 34 *High Risk*. Sementara, untuk individu dengan *chest pain* kategori *Typical Angina*, perlu memperhatikan beberapa atribut seperti Caa, Oldpeak, Sex, dan Exng (*Exercise Angina*). Setiap atribut pada rule di atas memiliki pengaruh yang cukup signifikan terhadap risiko penyakit serangan jantung bagi individu. Setelah melalui proses dalam membangun dan menguji model, maka dilakukan evaluasi kinerja model dengan menggunakan *confusion matrix* seperti pada Tabel 4 berikut:

Tabel 4. Confusion Matrix

	true High Risk	true Low Risk	class Precision
pred. High Risk	154 (TP)	38 (FP)	80.21%
pred. Low Risk	11 (FN)	100 (TN)	90.09%
class Recall	93.33%	72.46%	
accuracy: 83.98% +/- 18.91% (micro average: 83.83%)			

Berdasarkan *confusion matrix* yang tercantum pada tabel di atas, terlihat bahwa terdapat 154 *case* yang diprediksi mengalami High Risk serangan jantung dan sesuai dengan kenyataannya (TP). Terdapat 38 *case* diprediksi mengalami High Risk, tetapi kenyataannya mengalami Low Risk (FP).

11 *case* diprediksi mengalami Low Risk, tetapi kenyataannya mengalami High Risk (FN), dan terdapat 100 *case* diprediksi mengalami Low Risk yang sesuai dengan kenyataannya (TN).

Adapun metrik *Accuracy*, *Precision*, *Recall*, dan *Specificity* digunakan sebagai indikator untuk mengevaluasi kinerja model yang dihasilkan (Arkamil and Jambak, 2023). Berikut merupakan hasil perhitungan dari

beberapa indikator evaluasi kinerja model yang disebutkan:

a. Accuracy

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100\%$$

$$= \frac{154 + 100}{154 + 11 + 38 + 100} \times 100\%$$

$$= \frac{254}{303} \times 100\%$$

$$Accuracy = 83,83\%$$

b. Precision

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$= \frac{154}{154 + 38} \times 100\%$$

$$= \frac{154}{192} \times 100\% = 80.21\%$$

c. Recall

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

$$= \frac{154}{154 + 11} \times 100\%$$

$$= \frac{154}{165} \times 100\% = 93.33\%$$

d. Specificity

$$Specificity = \frac{TN}{FP + TN} \times 100\%$$

$$= \frac{100}{38 + 100} \times 100\%$$

$$= \frac{100}{138} \times 100\% = 72.46\%$$

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil dari penelitian ini, dapat diketahui dari dataset yang ada berjumlah 303 data dengan tipe data numerik dan nominal dengan proses pengolahan data menggunakan algoritma C4.5 menghasilkan nilai akurasi terbaik sebesar 83,98% melalui kombinasi beberapa parameter dari proses Optimize Parameters (Grid). Diketahui bahwa faktor-faktor, seperti Cp (Chest Pain), Caa (n

Major Vessels), Oldpeak, Sex, dan Exng (Exercise Angina) memiliki kontribusi yang cukup signifikan dalam menentukan risiko penyakit serangan jantung. Selain itu, juga diperoleh nilai akurasi dari indikator evaluasi kinerja model dengan akurasi sebesar 83,83%, nilai precision sebesar 80,21%, nilai recall sebesar 93,33%, dan nilai specificity sebesar 72.46%. Dengan demikian, dapat disimpulkan bahwa hasil pengolahan data ini menunjukkan bahwa algoritma C4.5 cukup baik untuk digunakan dalam mengklasifikasi risiko penyakit serangan jantung serta mengetahui faktor-faktor yang menjadi penentu kriteria seseorang berisiko mengalami penyakit serangan jantung.

4.2. Saran

Pada penelitian selanjutnya diharapkan agar dapat menambahkan dataset untuk meningkatkan akurasi dari model yang dihasilkan. Selanjutnya dapat dilakukan optimalisasi untuk meningkatkan nilai akurasi.

5. DAFTAR PUSTAKA

- Agus Oka Gunawan, I.M. *et al.* (2023) 'Klasifikasi Penyakit Jantung Menggunakan Algoritma Decision Tree Series C4.5 Dengan Rapidminer', *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(2), pp. 73–83. Available at: <https://doi.org/10.47233/jteksis.v5i2.775>.
- Arkamil, R.F. and Jambak, M.I. (2023) 'Klasifikasi Tindakan Persalinan Pada Pasien Ibu Bersalin Menggunakan Metode Decision Tree C4. 5', *KLIK: Kajian Ilmiah Informatika dan Komputer*, 4(1), pp. 513–523.
- Cardiovascular diseases* (no date) *World Health Organization*. Available at: <https://www.who.int/health-topics/cardiovascular-diseases#> (Accessed: 31 March 2024).
- Damanik, S.D. and Jambak, M.I. (2023) 'Klasifikasi Customer Churn pada Telekomunikasi Industri Untuk Retensi Pelanggan Menggunakan Algoritma C4. 5', *KLIK: Kajian Ilmiah Informatika dan Komputer*, 3(6), pp. 1303–1309.
- Han, J., Kamber, M. and Pei, J. (2012) *Data Mining: Concepts and Techniques Third Edition*. UK: Elsevier.
- Heart attack (myocardial infarction)* (2023) *Harvard Health Publishing*. Available at: https://www.health.harvard.edu/a_to_z/heart-attack-myocardial-infarction-a-to-z (Accessed: 31 March 2024).
- Maulana, Y., Winanjaya, R. and Rizki, F. (2022) 'Penerapan Data Mining dengan Algoritma C. 45 Dalam Memprediksi Penjualan Tempe', *Bulletin of Computer Science Research*, 2(2), pp. 53–58.
- North, M. (2016) *Data Mining for the Masses*. CreateSpace Independent Publishing Platform. Available at: <https://books.google.co.id/books?id=LqCfjwEACAAJ>.
- Santoso, M. *et al.* (2023) 'Klasifikasi Potensi Penyakit Jantung Menggunakan Algoritma C4. 5', *Jurnal INSAN Journal of Information System Management Innovation*, 3(2), pp. 96–103.
- Sepharni, A., Hendrawan, I.E. and Rozikin, C. (2022) 'Klasifikasi Penyakit Jantung dengan Menggunakan Algoritma C4. 5', *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 7(2), pp. 117–126.
- Suherman, S. *et al.* (2024) 'KLASIFIKASI GAYA BELAJAR PADA MAHASISWA MENGGUNAKAN METODE ALGORITMA C4. 5', *JIKA (Jurnal Informatika)*, 8(1), pp. 104–111.
- Tasia, E. *et al.* (2023) *Klasifikasi Penyakit Gagal Jantung Menggunakan Supervised Learning*. Available at: <https://journal.irpi.or.id/index.php/senti mas>.
- What Is a Heart Attack?* (2022) *National Heart, Lung, and Blood Institute*. Available at: <https://www.nhlbi.nih.gov/health/heart-attack> (Accessed: 31 March 2024).