



Comparison of C4.5 and Naive Bayes Classification Algorithms for Predicting Heart Failure

Perbandingan Algoritma Klasifikasi C4.5 Dan Naive Bayes untuk Memprediksi Gagal Jantung

Tiara Afrah Afifah¹, Refri Martiansah², Mohd Alviyoni³, Abdullah Arifin⁴

^{1,2,3,4}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
UIN Sultan Syarif Kasim Riau, Indonesia

E-Mail :¹tiaraafrah277@gmail.com, ²refrimartian@gmail.com,
³alviyoniexxon@gmail.com, ⁴arifin907827@gmail.com

Corresponding Author: Tiara Afrah Afifah

Abstract

Cardiovascular disease is a non-communicable disease which is the number one cause of death every year. One of the cardiovascular diseases with increasing incidence and prevalence is heart failure. Heart failure is a complex clinical syndrome characterized by a decrease in the heart's ability to pump blood to meet the body's need for oxygen. Classification is one of the techniques of data mining that can be used in the prediction process. Classification can be processed using a Decision Tree, namely the C4.5 and Naive Bayes algorithms. This study aims to classify and apply data mining classification. The results of data classification were evaluated using the Confusion Matrix and the ROC curve to determine the level of accuracy using the Decision Tree C4.5 algorithm which was 86.78% and the AUC value was 0.867, while the accuracy value for Naive Bayes was 89.18% with the AUC value being 0.947. so it can be said that the Naive Bayes algorithm has good predictive results in predicting heart failure in a person.

Keywords: Classification, C4.5, Decision Tree, Heart Failure, Naive Bayes.

Abstrak

Penyakit kardiovaskuler merupakan penyakit tidak menular yang menjadi penyebab kematian nomor satu setiap tahunnya. Salah satu penyakit kardiovaskuler dengan insiden dan prevalensi yang semakin meningkat yaitu gagal jantung. Gagal jantung merupakan sindrom klinis yang bersifat kompleks dengan karakteristik penurunan kemampuan jantung dalam memompa darah untuk memenuhi kebutuhan oksigen dalam tubuh. Klasifikasi merupakan salah satu teknik dari data mining yang bisa digunakan dalam proses prediksi. Klasifikasi bisa di proses dengan Decision Tree yaitu dengan algoritma C4.5 dan Naive Bayes. Penelitian ini bertujuan untuk melakukan klasifikasi serta menerapkan klasifikasi data mining. Hasil klasifikasi data di evaluasi dengan menggunakan Confusion Matrix dan kurva ROC untuk mengetahui tingkat hasil akurasi dengan menggunakan algoritma Decision Tree C4.5 adalah 86.78% dan nilai AUC adalah 0.867, sedangkan nilai akurasi pada Naive Bayes adalah 89.18% dengan nilai AUC adalah 0.947. sehingga dapat dikatakan bahwa algoritma Naive Bayes memiliki hasil prediksi yang baik dalam memprediksi penyakit Gagal Jantung pada seseorang.

Kata Kunci: C4.5, Decision Tree, Gagal Jantung, Klasifikasi, Naive Bayes.

1. PENDAHULUAN

Kardiovaskuler ialah penyakit yang tidak menular akan tetapi menjadi salah satu penyebab kematian terbesar pertama setiap tahunnya. Diantara banyaknya penyakit kardiovaskuler dengan kejadian serta karakteristik yang semakin tinggi adalah gagal jantung[1]. Penyakit ini termasuk sindrom klinis yang bersifat

komplek dengan ciri menurunnya kemampuan jantung seseorang dalam memompa darah guna mencukupi oksigen pada tubuh seorang.[2].

Banyaknya yang terkena kardiovaskuler jantung pada Asia Tenggara hingga 3 kali lebih besar dibandingkan Amerika dan Eropa ialah sebanyak 4,5 – 6,7%: 0,5 – 2%. Pada Indonesia, berdasarkan data Riset Kesehatan di tahun 2013 menyampaikan proporsi gagal jantung dengan angka 0,3% diprediksi kurang lebih 530.060 orang. Meskipun manajemen farmakologis serta medis sudah meningkat dengan pesat, akan tetapi angka kematian yang ditimbulkan gagal jantung tetap tinggi yaitu mencapai 50% dalam 5 tahun semenjak diagnosa ditegakkan[3].

Teknik analisa konvensional secara manual tidak lagi efektif untuk digunakan dalam mendiagnosa karena teknik ini dapat menyebabkan bias, kesalahan dan biaya medis yang berlebihan yang berpengaruh terhadap kualitas layanan yang diberikan kepada pasien. Semakin berkembangnya sistem berpedoman pada wawasan medis, kebutuhan untuk menggunakan sistem komputer untuk kemampuan analisis didalam menentukan penyakit semakin penting. Sebab itu, dikembangkan sistem pengetahuan yang terkomputerisasi, sehingga efektif dan efisien untuk mendiagnosa penyakit[4].

Oleh karena itu, penelitian ini dilakukan dengan tujuan untuk membantu menyelesaikan permasalahan tersebut dengan data mining yang berfungsi untuk memprediksi gagal jantung. Pada penelitian ini diperlukan suatu metode atau teknik yang dapat mengolah data-data yang sudah ada. Salah satu metode yang digunakan adalah teknik data mining. Penggunaan data mining dengan algoritma C4.5 dan Naive Bayes sebagai pilihan untuk memprediksi gagal jantung dapat menjadi alternatif pilihan yang tepat, tetapi sampai saat ini belum ada penelitian yang menjelaskan algoritma apa yang paling akurat untuk memprediksi gagal jantung[5]. Pada penelitian ini akan dilakukan komparasi data mining algoritma C4.5 dan Naive Bayes untuk mengetahui algoritma yang memiliki akurasi yang lebih tinggi dalam memprediksi gagal jantung.

2. BAHAN DAN METODE

2.1. Kajian Literatur

Menurut (Cahya Putri Buani, 2021) dalam penelitiannya yang berjudul “Penerapan Algoritma Naive Bayes dengan Seleksi Fitur Algoritma Genetika Untuk Prediksi Gagal Jantung”, metode yang digunakan pada penelitiannya adalah dengan dilakukannya peningkatan akurasi dari naive bayes dengan menggunakan algoritma genetika. Algoritma Genetika digunakan untuk mereduksi attribute-attribut yang kurang dominan karena metode algoritma genetika memiliki kemampuan dalam menseleksi atribut yang kurang dominan. Dalam penelitiannya Metode Naive Bayes menggunakan algoritma genetika dapat memberikan kemudahan dalam menghitung dan dapat menentukan kemungkinan-kemungkinan yang akan terjadi, serta naive bayes dapat menghitung dalam jumlah data yang besar. Namun pada metode ini terdapat kekurangan yang dapat berpengaruh pada akurasi yang disebabkan oleh fitur Naive Bayes yang tidak selalu dapat diterapkan ada penelitian yang menjelaskan algoritma apa yang paling akurat untuk memprediksi gagal jantung[6].

Menurut (Rohman & Rochcham, 2018) dalam penelitiannya yang berjudul “Model Algoritma C4.5 untuk Prediksi Penyakit Jantung” menjelaskan bahwa metode Decision Tree mempunyai kelebihan-kelebihan dalam mengolah dataset penyakit jantung yaitu dari segi; kecepatan dalam klasifikasi, tiap atribut bersifat diskrit, binari dan continue, serta transparansi pengetahuan atau klasifikasi sehingga dalam penelitiannya memilih metode Decision Tree atau C4.5 dalam memprediksi penyakit jantung sehingga terbentuk modelnya, dengan mengoptimalkan atribut-attribut yang berasal dari dataset yang terpercaya untuk memprediksi penyakit jantung dengan tujuan agar akurasi menjadi meningkat ada penelitian yang menjelaskan algoritma apa yang paling akurat untuk memprediksi gagal jantung[7].

2.2. Data Mining

Data Mining merupakan salah satu cara untuk menentukan alur serta pandangan atraktif pada data yang berjumlah besar[8]. Data merupakan bagian dari Knowledge Discovery Data (KDD) yaitu proses untuk meringkas sebuah informasi yang bermanfaat yang sebelumnya belum diketahui dan tidak terlihat[9]. Penggunaan teknik data mining diharapkan dapat membantu dalam mempercepat proses pengambilan keputusan[10]. Knowledge discovery sebagai suatu proses terdiri atas pembersihan data (data cleaning), integrasi data (data integration), pemilihan data (data selection), transformasi data (data transformation), data mining, evaluasi pola (pattern evaluation) dan penyajian pengetahuan (knowledge presentation)[11].

2.3. Klasifikasi

Klasifikasi merupakan salah satu teknik data mining yang bertujuan untuk menggambarkan suatu model pengetahuan berdasarkan ciri persamaan atau perbedaan dari suatu kelas data[12]. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa

algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Artificial Neural Networks[13].

2.4. C4.5

Metode Algoritma C4.5 merupakan algoritma yang digunakan untuk membuat decision tree atau pohon keputusan. Decision tree merupakan salah satu proses klasifikasi dan prediksi yang sangat kuat dan terkenal[14]. Algoritma C4.5 sebagai klasifikasi yang paling sederhana dan mudah diimplementasikan, akan tetapi Algoritma C4.5 masih memiliki kelemahan untuk menangani data berdimensi tinggi[18]. Dalam menggunakan C4.5 harus dilengkapi dengan variabel atau atribut pada data yang digunakan[15].

2.5. Naïve Bayes

Metode Pengklasifikasi Naive Bayes adalah pengklasifikasi paling sederhana dan paling umum digunakan [16]. Naive Bayes merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema Bayes (aturan bayes) dengan asumsi independensi (ketidak tergantungan) yang kuat (naif) [10]. Pada penelitian ini akan melakukan perbandingan data mining algoritma C4.5 dan Naive Bayes untuk mengetahui algoritma yang memiliki akurasi yang lebih tinggi dalam mendeteksi penyakit jantung.

2.6 Evaluasi Dan Validasi

Evaluasi dan Validasi yang di lakukan dalam penelitian ini menggunakan model Confusion Matrix dan Kurva ROC.

Confusion Matrix

Matriks konfusi ialah cara yang dapat digunakan untuk mengukur performa metode klasifikasi dalam memprediksi kebenaran ataupun kesalah suatu objek. Matriks konfusi dapat membandingkan dampak dari kelas input awal yang berisi informasi nilai asli dan nilai perkiraan pada hasil klasifikasi[17]. Confusion matrix memuat detail klasifikasi, dengan class yang diprediksi [5].

Kurva ROC

Kurva ROC dapat digunakan untuk memberikan perbandingan pengelompokan dengan cara visual. Kurva ini merupakan pengekspresian dari Confusion Matrix, Kurva ini memiliki grafik 2D dengan Positif false adalah garis horizontal dan positif true adalah garis vertikal. ROC (Receiver Operating Characteristic) adalah suatu metode yang dapat dipakai untuk menggambarkan akurasi diferensial suatu tes diagnostik untuk menentukan apakah seseorang mengidap penyakit tertentu.

Keakuratan kinerja AUC dapat dibagi menjadi lima kategori, yaitu:

1. Angka Akurasi Atau Ketepatan 0,90 – 1,00 = Klasifikasi Sangat Baik
2. Angka Akurasi Atau Ketepatan 0,80 – 0,90 = Klasifikasi Baik
3. Angka Akurasi Atau Ketepatan 0,70 – 0,80 = Klasifikasi Wajar
4. Angka Akurasi Atau Ketepatan 0,60 – 0,70 = Klasifikasi Buruk
5. Angka Akurasi Atau Ketepatan 0.50 – 0.60 = Gagal

2.7. Metode Penelitian

Kerangka Pemikiran

Pada penelitian ini digunakan bentuk penelitian menggunakan model penelitian eksperimen. Penelitian ini bertujuan untuk melakukan perbandingan dan evaluasi pada algoritma klasifikasi data mining. Studi eksperimen ini menggarisbawahi teori-teori pada penelitian terdahulu. Penelitian ini menggunakan eksperimen komparatif berdasarkan framework pada gambar 1.

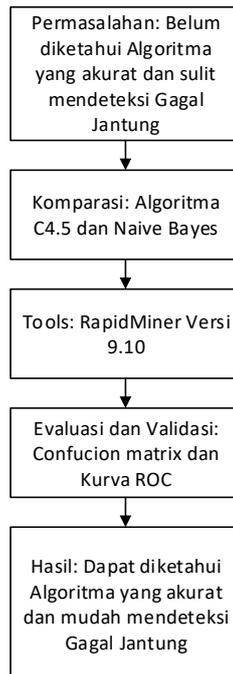
Tahap awal pada kerangka pemikiran penelitian ini adalah analisis masalah penyakit jantung, selanjutnya mengusulkan metode (model) untuk menyelesaikan masalah tersebut berupa algoritma C4.5 dan Naive Bayes. Alat yang dimanfaatkan yaitu RapidMiner versi 9.10, dataset didapatkan dari, <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>, uji evaluasi dan validasi menggunakan matriks konfusi dan kurva ROC untuk mengukur akurasi. Berdasarkan hasil yang diperoleh antar algoritma, diperoleh algoritma terbaik untuk memprediksi gagal jantung.

Langkah Penelitian

Pada penelitian ini, penulis menggunakan sistem komputer HP Notebook - 14-am015tx dengan konfigurasi hardware yang terdiri dari Processor Intel®Core™ i5-6200U CPU @ 2.30GHz (4 CPUs), ~2.4GHz, RAM 4 GB, SSD 224 GB, dan software yang terdiri dari Sistem operasi Windows 10 Pro dan tools data mining Rapid Miner. Penelitian ini dilakukan dengan menjalankan beberapa langkah proses penelitian yaitu:

1. Pengumpulan Data
2. Pengolahan data awal

3. Pengukuran penelitian
4. Analisa komparasi hasil



Gambar 1. Kerangka Pemikiran

3. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini bersumber dari alamat web: <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> yang merupakan dataset gagal jantung hasil pemeriksaan terhadap 500 orang yang terdiri dari 12 atribut, dengan 11 atribut predictor dan 1 atribut tujuan. Data dapat dilihat pada tabel 1.

Table 1. Data Pasien Gagal Jantung

No	Age	Sex	CPT	RBP	Cholesterol	FBS	RECG	MaxHR	EA	Oldpeak	ST_S	HeartDisease
1	40	M	ATA	140	289	0	Normal	172	N	0	Up	0
2	49	F	NAP	160	180	0	Normal	156	N	1	Flat	1
3	37	M	ATA	130	283	0	ST	98	N	0	Up	0
4	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
5	54	M	NAP	150	195	0	Normal	122	N	0	Up	0
6	39	M	NAP	120	339	0	Normal	170	N	0	Up	0
7	45	F	ATA	130	237	0	Normal	170	N	0	Up	0
8	54	M	ATA	110	208	0	Normal	142	N	0	Up	0
9	37	M	ASY	140	207	0	Normal	130	Y	1.5	Flat	1
10	48	F	ATA	120	284	0	Normal	120	N	0	Up	0
...
915	45	M	TA	110	264	0	Normal	132	N	1.2	Flat	1
916	68	M	ASY	144	193	1	Normal	141	N	3.4	Flat	1
917	57	M	ASY	130	131	0	Normal	115	Y	1.2	Flat	1
918	57	F	ATA	130	236	0	LVH	174	N	0	Flat	1
919	38	M	NAP	138	175	0	Normal	173	N	0	Up	0

Tabel 1 diatas merupakan sample dari data penyakit jantung dengan keterangan yang bisa dilihat di tabel 2.

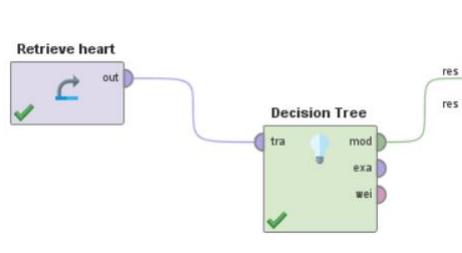
Table 2. Keterangan Data Penyakit Jantung

Atribut	Keterangan	Tipe	Nilai Atribut
Age	Usia	Integer	Angka
Sex	Jenis Kelamin	Binominal	Huruf
Chest pain type	Jenis nyeri dada	Polynominal	Huruf
Resting BP(blood pressure)	Tekanan darah	Integer	Angka

Atribut	Keterangan	Tipe	Nilai Atribut
Cholesterol	Kolesterol	Integer	Angka
Fasting BS(blood sugar)	Gula darah	Binominal	Angka
Resting ECG	Elektrokardiogram	Polynomial	Huruf
Max HR(heart rate)	Denyut jantung maksimal	Integer	Angka
Exercise Angina	Angin duduk	Binominal	Huruf
Oldpeak	Depresi	Real	Angka
ST_Slope	Kemiringan ST	Polynomial	Huruf
Heart Disease	Penyakit jantung	Binominal	Angka

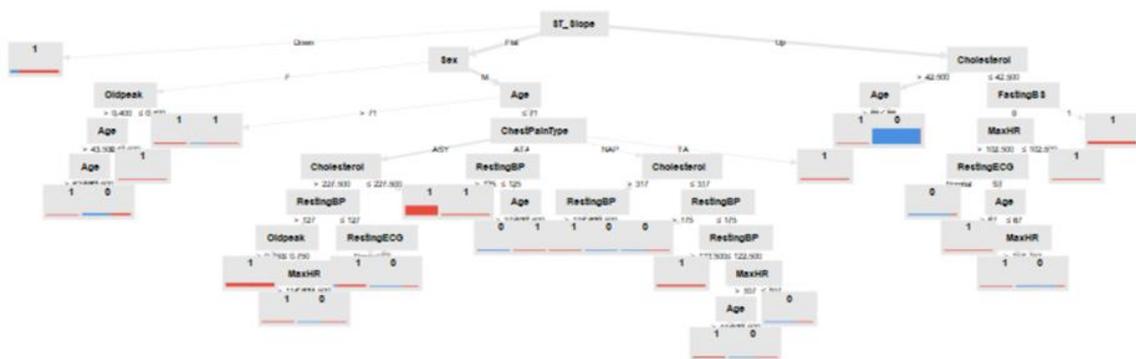
3.1. Pengolahan Data Awal

Tahap awal pada pengolahan data ini dimulai dengan menetapkan data. Berdasarkan data pada tabel 1 akan dibentuk suatu bentuk descision tree dengan tools RapidMiner 9.10. adapun bentuk model pohon keputusan atau descision tree nya adalah sebagai berikut:



Gambar 2. Desain Model Algoritma C4.5

Gambar 2 diatas merupakan gambaran model bentuk dari data yang di dapatkan yaitu data penyakit jantung yang selanjutnya akan direlasikan dalam Algoritma C4.5 dengan memanfaatkan tools RapidMiner versi 9.10. Dari desain tersebut akan menghasilkan desain sebuah pohon keputusan yang bisa terlihat pada gambar 3.



Gambar 3. Hasil Klasifikasi Menggunakan Algoritma C4.5

3.2 Pengukuran Penelitian

a. Confusion Matrix

Hasil *Confusion matrix* berdasarkan pemodelan hasil olahan data dengan algoritma C4.5 didapatkan hasil akurasi sebesar 86.78% dapat dilihat di tabel 3.

Tabel 3. Hasil *Confusion Matrix* C4.5

	True Positif	True Negatif	Class Precision
Pred. Positif	189	39	82.89%
Pred. Negatif	27	244	90.04%
Class Recall	87.50%	86.22%	

Berdasarkan tabel 3 menunjukkan bahwa, dari 500 data pasien penyakit jantung yang di uji, terdapat 216 pasien yang tidak terdeteksi gagal jantung namun berdasarkan hasil dari tabel confusion matrix diatas terdapat 189 pasien yang diperikarakan negatif, sehingga hasilnya benar yaitu negatif, 39 pasien diperkirakan negatif

namun hasilnya positif, dan 283 pasien yang diperkirakan positif namun pada tabel diatas menampilkan terdapat 27 pasien yang diperkirakan positif namun hasilnya negatif dan 244 pasien yang diperkirakan positif dengan hasil benar yaitu positif sesuai dengan prediksi. Adapun nilai hasil ketepatan dengan Algoritma C4.5 ialah 86.78%, dan untuk mendapatkan nilai akurasinya dapat dihitung sebagai berikut:

Keterangan:

$$TP = 244 \quad TN = 189$$

$$FP = 39 \quad FN = 27$$

$$\begin{aligned} \text{Akurasi} &= (TP+TN) / (TP+TN+FP+FN) \\ &= (244+189) / (244+189+39+27) \\ &= 0.8678 \text{ (86.78\%)} \end{aligned}$$

Sedangkan hasil *Confusion Matrix* untuk *Naive Bayes* yang berakurasi dengan angka yang didapat ialah 89.18% bisa dilihat di tabel 4.

Tabel 4. Hasil *Confusion Matrix* Naïve Bayes

	True Positif	True Negatif	Class Precision
Pred. Positif	182	20	90.10%
Pred. Negatif	34	263	88.55%
Class Recall	84.26%	92.93%	

Berdasarkan tabel 4 menunjukkan bahwa, dari 500 data pasien penyakit jantung yang di uji, terdapat 216 pasien yang tidak terdeteksi gagal jantung namun berdasarkan hasil dari tabel confusion matrix diatas terdapat 182 pasien yang diperkirakan negatif, sehingga hasilnya benar yaitu negatif, 20 pasien diperkirakan negatif namun hasilnya positif, dan 283 pasien yang diperkirakan positif namun pada tabel diatas menampilkan terdapat 34 pasien yang diperkirakan positif namun hasilnya negatif dan 263 pasien yang diperkirakan positif dengan hasil benar yaitu positif sesuai dengan prediksi. Adapun nilai hasil ketepatan dengan Algoritma Naïve Bayes ialah 89.18%, dan untuk mendapatkan angka akurasinya dapat dihitung sebagai berikut:

Keterangan:

$$TP = 263 \quad TN = 182$$

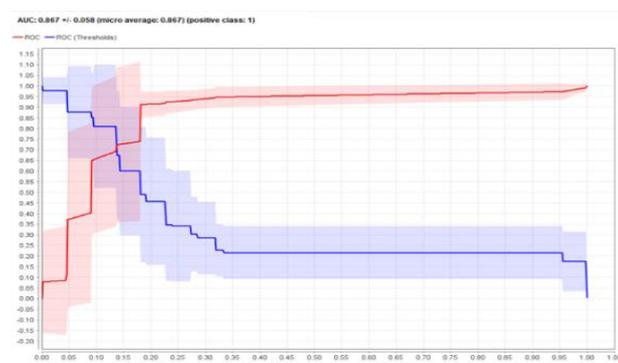
$$FP = 20 \quad FN = 34$$

$$\begin{aligned} \text{Akurasi} &= (TP+TN) / (TP+TN+FP+FN) \\ &= (263+182) / (263+182+20+34) \\ &= 0.8918 \text{ (89.18\%)} \end{aligned}$$

Berdasarkan hasil pengujian menggunakan confusion matrix diketahui bahwa model C4.5 memiliki nilai akurasi sebesar 86.78%. Sedangkan model Naive Bayes memiliki nilai akurasi sebesar 89.18%. Berdasarkan pengujian tersebut dapat disimpulkan bahwa akurasi pada Naïve Bayes lebih tinggi dibandingkan dengan algoritma C4.5 dengan selisih 2.40 %.

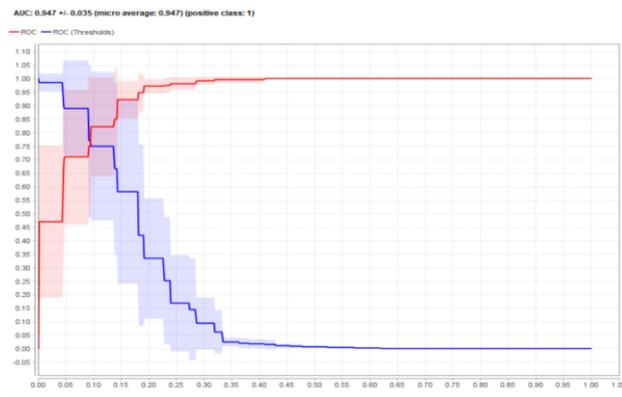
b. Kurva ROC

Pengujian pada dataset akan dilakukan penilaian berdasarkan hasil perkiraan dalam bentuk grafik ROC pada C 4.5 dan naive bayes. Adapun visualisasi dari grafik ROC pada algoritma C4.5 seperti berikut.



Gambar 4. Grafik ROC dari Model Algoritma C4.5

Pada gambar 4 diatas merupakan angka tingkat positif false yang ditampilkan menggunakan garis vertikal yang memiliki angka 0 sampai 1,05 dan nilai tingkat positif true yang ditampilkan menggunakan garis horizontal yang memiliki 0 hingga 1,15. Pada gambar diatas menghasilkan ROC dan AUC yaitu 0.867 dengan hasil uji coba termasuk Klasifikasi Baik. Dibandingkan dengan visualisasi ROC menggunakan Naive Bayes dapat dilihat pada gambar 5.



Gambar 5. Grafik ROC dari Model Naive Bayes

Pada gambar 5 merupakan angka tingkat positif false yang ditampilkan menggunakan garis vertikal yang memiliki angka 0 sampai 1,05 dan nilai tingkat positif true yang ditampilkan menggunakan garis horizontal yang memiliki 0 hingga 1,10. Pada gambar diatas menghasilkan ROC dan AUC yaitu 0.947 dengan hasil uji coba termasuk Klasifikasi Sangat Baik.

3.3 Analisa Hasil Komparasi

Berdasarkan hasil pengujian yang telah dilakukan dengan matriks konfusi dan ROC maka bentuk klasifikasi C4.5 dan Naïve Bayes yang dapat dilihat pada tabel 5 berikut.

Tabel 5. Perbandingan Hasil Akurasi dan Nilai AUC

	Accuracy	AUC
Decision Tree	86.78%	0.867
Naïve Bayes	89.18%	0.947

Pada tabel 5 diatas dapat dilihat bahwa evaluasi dengan memanfaatkan kurva ROC sebagai bentuk Decision Tree dengan hasil AUC yaitu 0.867 dengan nilai diagnosa Klasifikasi Baik dan akurasi dengan nilai 86.78%. Sedangkan untuk algoritma Naive Bayes mendapatkan nilai AUC yaitu 0.947 dengan nilai pengujian Klasifikasi Sangat Baik dan hasil akurasi dengan nilai 89.18%. Adapun selisih nilai keduanya sebesar 2.40%. Dengan demikian bisa diambil kesimpulan berupa Algoritma Naive Bayes ialah algoritma yang dapat dimanfaatkan dalam proses prediksi untuk menyelesaikan permasalahan dalam memprediksi gagal jantung.

4. KESIMPULAN

Setelah dilakukan pengujian dan analisis dalam memprediksi gagal jantung dengan memanfaatkan model C4.5 dan Naive Bayes dengan tujuan guna mendapati algoritma mana yang memiliki angka akurasi paling besar, maka dapat disimpulkan bahwa, berdasarkan uji hasil banding C4.5 dan Naive Bayes setelah diuji ketepatannya dengan memanfaatkan Confusion Matrix dan Kurva ROC maka diketahui nilai akurasi C4.5 ialah 86.78% dengan nilai AUC 0.867, adapun nilai akurasi pada Naive Bayes ialah 89.18% dengan nilai AUC 0.947. Berdasarkan hasil perbandingan pada kedua algoritma tersebut dapat diambil kesimpulan bahwa tingkat akurasi pada model Naive Bayes lebih tinggi, dengan selisih tingkat akurasi sebesar 2.40% serta peningkatan nilai AUC sebesar 0.080 dengan hasil diagnosa perikiraan termasuk Klasifikasi Sangat Baik.

REFERENSI

- [1] S. Saida, H. Haryati, and L. Rangki, "Kualitas Hidup Penderita Gagal Jantung Kongestif Berdasarkan Derajat Kemampuan Fisik dan Durasi Penyakit," *Faletehan Heal. J.*, vol. 7, no. 02, pp. 70–76, 2020, doi: 10.33746/fhj.v7i02.134.
- [2] P. Gagal, J. Di, and R. S. D. Mangusada, "ISSN 2089-4503 (cetak)," vol. 8487, no. 1, pp. 39–47, 2020.
- [3] D. Prihatiningsih and T. Sudyasih, "Perawatan Diri Pada Pasien Gagal Jantung," *J. Pendidik. Keperawatan Indones.*, vol. 4, no. 2, 2018, doi: 10.17509/jpki.v4i2.13443.
- [4] Suparyanto dan Rosad (2015, "濟無No Title No Title No Title," *Suparyanto dan Rosad (2015, vol. 5,*

- no. 3, pp. 248–253, 2020.
- [5] Fatmawati, “Perbandingan Algoritma Klasifikasi Data Mining Model C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Diabetes,” *J. Techno Nusa Mandiri*, vol. XIII, no. 1, p. 50, 2016.
- [6] D. Cahya Putri Buani, “Penerapan Algoritma Naïve Bayes dengan Seleksi Fitur Algoritma Genetika Untuk Prediksi Gagal Jantung,” *EVOLUSI J. Sains dan Manaj.*, vol. 9, no. 2, pp. 43–48, 2021, doi: 10.31294/evolusi.v9i2.11141.
- [7] A. Rohman and M. Rochcham, “Model Algoritma C4.5 Untuk Prediksi Penyakit Jantung,” *Neo Tek.*, vol. 4, no. 2, pp. 52–55, 2018, doi: 10.37760/neoteknika.v4i2.1228.
- [8] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, “Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review,” *Fakt. Exacta*, vol. 13, no. 1, p. 35, 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [9] C. Agustina, “Analisa Nasabah Potensial Tabungan Deposito Berjangka Menggunakan Teknik Klasifikasi Data Mining,” *J. Teknol. Inf. dan Terap.*, vol. 5, no. 2, pp. 105–112, 2019, doi: 10.25047/jtit.v5i2.88.
- [10] H. D. Wijaya and S. Dwiasnati, “Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat,” *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.
- [11] D. Firdaus, “Penggunaan Data mining dalam kegiatan pembelajaran,” vol. 6, no. 2, pp. 91–97, 2017.
- [12] T. Novika, P. Poningsih, H. Okprana, A. P. Windarto, and H. Siahaan, “Penerapan Data Mining Klasifikasi Tingkat Pemahaman Siswa Pada Pelajaran Matematika,” *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 9, 2021, doi: 10.30865/mib.v5i1.2498.
- [13] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, “Metode-metode Klasifikasi,” *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
- [14] R. Girsang, E. F. Ginting, and M. Hutasuhut, “Penerapan Algoritma C4.5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah,” *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 449, 2022, doi: 10.53513/jursi.v1i4.5727.
- [15] A. K. Lalo, P. Batarius, and Y. C. H. Siki, “Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia,” *J. Tek. Inform. UNIKA St. Thomas*, no. June, pp. 1–12, 2021, doi: 10.54367/jtiust.v6i1.1089.
- [16] W. Yulita *et al.*, “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier,” *Jdmsi*, vol. 2, no. 2, pp. 1–9, 2021.
- [17] A. Harun and D. P. Ananda, “Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decision Tree,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 58–63, 2021.
- [18] D. Saputra, W. Irmayani, D. Purwaningtias, and J. Sidauruk, “A Comparative Analysis of C4.5 Classification Algorithm, Naïve Bayes and Support Vector Machine Based on Particle Swarm Optimization (PSO) for Heart Disease Prediction,” *Int. J. Adv. Data Inf. Syst.*, vol. 2, no. 2, pp. 84–95, 2021, doi: 10.25008/ijadis.v2i2.1221.