

Komparasi Algoritma Data Mining dan Perancangan Aplikasi Prediksi Harapan Hidup Pasien Gagal Jantung

Comparison of Data Mining Algorithms and Design of Life Expectancy Prediction Applications in Heart Failure Patients

Agustiena Merdekawati^{a,1}

Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Jakarta^a
agustiena.atd@bsi.ac.id¹

ABSTRAK

Jantung merupakan organ tubuh yang paling vital. Penyakit gagal jantung merupakan penyebab kematian dengan jumlah kasus terbesar. Oleh karena itu, perlu adanya suatu perkiraan mengenai faktor terbesar harapan hidup pada penderita penyakit gagal jantung, sehingga dapat menurunkan angka kematian. Dalam memprediksi harapan hidup penyakit gagal jantung dengan menggunakan Knowledge Discovery in Database (KDD) dapat ditemukan pola prediksi harapan hidup penyakit gagal jantung, sehingga dapat menekan angka kematian. Dalam penelitian ini menggunakan algoritma C4.5 dan algoritma C4.5 dengan PSO (Particle Swarm Optimization) untuk mendapatkan pola prediksi harapan hidup penyakit gagal jantung yang selanjutnya didapatkan persentase precision, recall dan accuracy. Penelitian ini yaitu menghasilkan pola prediksi harapan hidup penyakit gagal jantung dengan Kriteria lama waktu tindakan memiliki prioritas utama. Dengan menggunakan algoritma C4.5 didapatkan akurasi sebesar 73.33%, sedangkan menggunakan algoritma C4.5 dan PSO didapatkan akurasi sebesar 99.00%, sehingga dapat disimpulkan berdasarkan tingkat akurasi bahwa pemodelan algoritma C4.5 dan PSO memiliki keakuratan yang lebih tinggi dibandingkan algoritma C4.5. Dengan menggunakan algoritma C4.5 didapatkan akurasi grafik ROC sebesar 0.897%, sedangkan menggunakan algoritma C4.5 dan PSO didapatkan akurasi grafik ROC sebesar 1.00%, sehingga dapat disimpulkan berdasarkan tingkat akurasi grafik ROC bahwa pemodelan algoritma C4.5 dan PSO memiliki keakuratan yang lebih tinggi dibandingkan algoritma C4.5. Oleh karena itu, dibuatlah rancangan aplikasi berdasarkan metode data mining yang tingkat akurasinya paling tinggi agar dapat digunakan untuk memprediksi harapan hidup pasien penyakit gagal jantung.

Kata Kunci : *Komparasi, Algoritma C4.5, algoritma C4.5 dengan PSO, Harapan hidup, Gagal Jantung*

ABSTRACT

The heart is the most vital organ of the body. Heart failure is the leading cause of death with the largest number of cases. Therefore, it is necessary to estimate the biggest factor in life expectancy in patients with heart failure, so as to reduce mortality. In predicting the life expectancy of heart failure by using Knowledge Discovery in Database (KDD) it is possible to find predictive patterns of life expectancy for heart failure, so that it can reduce mortality. In this study using the C4.5 algorithm and the C4.5 algorithm with PSO (Particle Swarm Optimization) to obtain a predictive pattern of life expectancy for heart failure which then obtained the percentage of precision, recall and accuracy. This research is to produce a predictive pattern of life expectancy for heart failure with the criteria for the length of time the action has a top priority. By using the C4.5 algorithm, an accuracy of 73.33% is obtained, while using the C4.5 and PSO algorithms an accuracy of 99.00% is obtained, so it can be concluded based on the accuracy level that the C4.5 and PSO algorithm modeling has a higher accuracy than the C4.5 algorithm. . By using the C4.5 algorithm, the ROC graph accuracy is 0.897%, while using the C4.5 and PSO algorithms the ROC graph accuracy is 1.00%, so it can be concluded based on the ROC graph accuracy level that the C4.5 and PSO algorithm modeling has more accuracy. higher than the C4.5 algorithm.

Keywords : *Comparison, Algorithm C4.5, Algorithm C4.5 with PSO, Life expectancy, Heart Failure*

Disubmit: 03 September 2022

Info Artikel :
Direview: 11 Oktober 2022

Diterima: 29 November 2022

Copyright © 2022 – CSRID Journal. All rights reserved.

1. PENDAHULUAN

Jantung merupakan organ tubuh yang paling vital yang berfungsi dalam mempompa darah agar kebutuhan oksigen dan nutrisi terpenuhi. Peredaran darah dapat terganggu jika jantung mengalami gangguan atau kerusakan. Beberapa jenis penyakit jantung antara lain: jantung koroner, kelainan irama jantung, jantung bawaan, kelainan katup atau klep jantung, dan gagal jantung.

Berdasarkan data Rikesdas menunjukkan tanda dari penyakit jantung yaitu seperti hipertensi pada tahun 2013 sebesar 25,8% meningkat menjadi 34,1% pada tahun 2018, stroke pada tahun 2013 sebesar 12,1 per mil meningkat menjadi 10,9 per mil pada tahun 2018, penyakit jantung koroner dari tahun 2013 sampai 2018 tetap berada 1,5%, penyakit ginjal kronis pada tahun 2013 sebesar 0,2% meningkat menjadi 0,38% pada tahun 2018 [1].

Jumlah kematian pada penyakit gagal jantung sebesar 27% dengan sekitar 3 sampai 20 per 100- rang menderita gagal jantung, angka dari penderita gagal jantung terus meningkat sesuai dengan bertambahnya usia, yaitu 100 per 1000 orang pada rentang usia diatas 60 tahun. Pada penelitian Framingham tahun 2000 angka kematian dari penyakit gagal jantung ini dalam 5 tahun terakhir sebesar 42% pada wanita dan 62% pada pria. Berdasarkan data setiap tahunnya penduduk Amerika terdapat 3 juta penduduk yang mengalami gagal jantung dan setiap tahunnya mengalami pertambahan 400.000 orang. Sedangkan di Indonesia, penyakit gagal jantung menduduki peringkat nomer satu dalam kematian, yang sebelumnya pada peringkat ketiga.[2]

Diambil dari data kesehatan [3], penyakit jantung merupakan jumlah kasus terbanyak yang dibiayai oleh BPJS Kesehatan, yaitu 11. 592. 990 kasus.

Berdasarkan penjelasan sebelumnya, kematian akibat penyakit jantung dan penderita penyakit jantung setiap tahunnya bertambah, sehingga penyakit jantung ini menjadi suatu penyakit yang berbahaya namun tidak menular. Oleh karena itu, perlu adanya suatu perkiraan mengenai faktor terbesar harapan hidup pada penderita penyakit gagal jantung, sehingga dapat menurunkan angka kematian. Prediksi tersebut dapat di jadikan suatu acuan untuk saran pengobatan, perubahan gaya hidup dan perilaku penderita penyakit gagal jantung, dan pola diet untuk penderita penyakit gagal jantung. Prediksi ini juga dapat digunakan oleh seorang dokter agar tidak terjadi keterlambatan dalam menangani pasien Melalui data rekam medis pasien penyakit gagal jantung yang dianalisis menghasilkan suatu prediksi harapan hidup pasien penderita penyakit gagal jantung.

Beberapa penelitian yang telah dilakukan mengenai prediksi harapan hidup pada penderita gagal jantung. Pada penelitian yang dilakukan [4] mengenai prediksi gagal jantung dengan menggunakan algoritma naive bayes dan seleksi fitur algoritma genetika, memberikan hasil yaitu penelitian dengan menggunakan naive bayes menghasilkan tingkat akurasi sebesar 69,90% dan penelitian setelah menggunakan fitur algoritma genetika tingkat akurasi naik menjadi 96,67%, dengan selisih peningkatan akurasi sebesar 27,07%. Pada penelitian selanjutnya mengenai prediksi harapan hidup penyakit gagal jantung menggunakan algoritma PSO pada naive bayes, memberikan hasil akurasi sebesar 92,67% dengan nilai *Area Under ROC* (AUC) sebesar 0,908 [5].

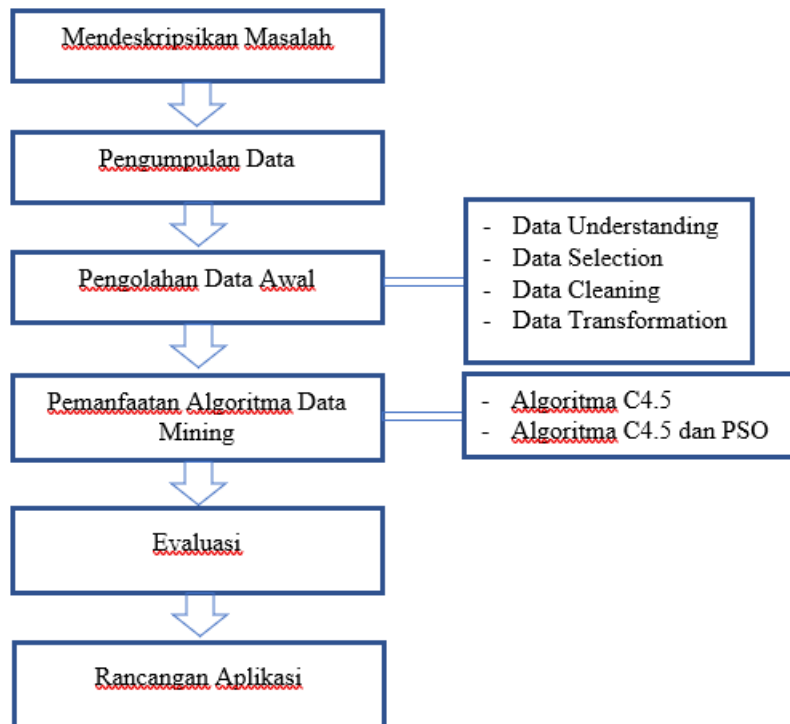
[6] Hasil penelitian dengan menggunakan algoritma C4.5 memberikan tingkat akurasi yang cukup tinggi. Sedangkan [4] dengan menambahkan algoritma PSO pada *naive bayes* memberikan hasil akurasi meningkat sebesar 96,67%.

Berdasarkan uraian latar belakang yang telah dibahas, dilakukan penelitian perbandingan prediksi harapan hidup penderita penyakit gagal jantung dengan menggunakan algoritma C4.5

dengan algoritma C4.5 dan PSO, serta merancang aplikasi prediksi harapan hidup penderita penyakit gagal jantung berdasarkan tingkat akurasi yang paling tinggi.

2. METODE PENELITIAN

Dalam melakukan penelitian ini membutuhkan beberapa tahapan yang digunakan agar peneliti mempunyai alur yang jelas sehingga menghasilkan penelitian yang baik. Berikut tahapan penelitian yang dilakukan sebagai berikut:



Gambar 1. Tahap Penelitian

2.1 Mendeskripsikan Masalah

Tahap awal penelitian ini dengan mendeskripsikan masalah yang terjadi pada penyakit gagal jantung. Pasien penderita penyakit gagal jantung tidak menyadari bahwa penyakit yang dideritanya merupakan penyakit yang dapat mengakibatkan kematian jika tidak ada pengendalian. Dengan adanya pengendalian maka dapat menurunkan angka kematian dari penyakit gagal jantung. Sehingga dengan adanya penelitian ini, pasien dapat memperbaiki perubahan gaya hidup dan perilaku penderita penyakit gagal jantung, dan pola diet dan dokter dapat memberikan informasi mengenai perubahan gaya hidup dan perilaku penderita penyakit gagal jantung, dan pola diet bagi pasien.

2.2 Pengumpulan Data

Setiap melakukan penelitian membutuhkan suatu dataset. Data yang digunakan berupa data sekunder yaitu *Heart Failure Clinical Records Dataset* yang diambil dari halaman situs *University of California Irvine Machine Learning Repository (UCI Repository)*. Data ini terdiri dari 299 data pasien yang mengalami gagal jantung, dimana memiliki 13 atribut klinis, yang terdiri dari 12 atribut *prededictor* dan 1 atribut tujuan. *Value data* dari masing-masing atribut dibuat dalam bentuk katagori. Berikut atribut yang menjadi parameter harapan hidup dari penyakit gagal jantung:

Tabel 1. Atribut data dan Nilai Data Katagori

No	Atribut	Nilai
1	Age (Usia Pasien)	40 - 50
		51- 60
		61-70
		71-80
		81-90
		91-95
2	Anaemia (Penurunan sel darah merah)	Ya
		Tidak
3	Creatine phosphokinase (CPK) (Tingkat enzim CPK)	26-192
		>192
		39-308
		>308
4	Diabetes (Memiliki diabetes)	Ya
		Tidak
5	Fraksi Ejeksi (Persentase darah yang meninggalkan jantung pada setiap kontraksi)	<55
		55-70
		>70
6	Hipertensi	Ya
		Tidak
7	Trombosit	<150000
		150000-450000
		>450000
8	Tingkat Kreatinin	<1.2
		>1.2
		<1.4
		>1.4
9	Tingkat natrium	<135
		135-145
		>145
10	Jenis Kelamin	P
		L
11	Merokok	Ya
		Tidak
12	Periode waktu tindakan	0-20
		21-40
		41-60
		61-80
		81-100
		101-120
		121-140
		141-160
		161-180
		181-200
		201-220
		221-240
		241-260
		261-280
281-290		

2.3 Pengolahan Data Awal

Pada tahap ini melakukan pendalaman dari data *Heart Failure Clinical*. Data yang sudah didapat dilakukan proses *understanding, selection, cleaning, dan transformatin data*.

1. Data Understanding

Data ini terdiri dari 299 data pasien yang mengalami gagal jantung. Data ini merupakan data asli yang diambil pada halaman situs *University of California Irvine Machine Learning Repository (UCI Repository)*. Berikut data aslinya sebagai berikut:

Tabel 2. Data Gagal Jantung Dari *UCI Repository*

age	anaemia	Kadar Kreatinin	diabetes	Fraksi Ejeksi	Tekanan Darah Tinggi	Trombosit	Serum Kreatinin	Serum Sodium	Jkel	Merokok	Waktu tindakan	Harapan hidup
75	0	582	0	20	1	265000	01:09	130	1	0	4	1
55	0	7861	0	38	0	263358.03	01:01	136	1	0	6	1
65	0	146	0	20	0	162000	01:03	129	1	1	7	1
50	1	111	0	20	0	210000	01:09	137	1	0	7	1
65	1	160	1	20	0	327000	02:07	116	0	0	8	1
90	1	47	0	40	1	204000	02:01	132	1	1	8	1
75	1	246	0	15	0	127000	01:02	137	1	0	10	1
60	1	315	1	60	0	454000	01:01	131	1	1	10	1
65	0	157	0	65	0	263358.03	01:05	138	0	0	10	1
80	1	123	0	35	1	388000	09:04	133	1	1	10	1
75	1	81	0	38	1	368000	4	131	1	1	10	1
62	0	231	0	25	1	253000	00:09	140	1	1	10	1
45	1	981	0	30	0	136000	01:01	137	1	0	11	1
50	1	168	0	38	1	276000	01:01	137	1	0	11	1

2. Data Selection (Pemilihan Data)

Data *selection* digunakan saat memproses data mining dengan penyimpanan data yang terpisah dengan penyimpanan data mining [7]. Proses data *selection* dengan memilih data yang tidak mempunyai hubungan dengan penelitian, sehingga atribut yang tidak diperlukan dalam penelitian akan dihapus. Pada data ini tidak terdapat atribut yang dihapus, dikarenakan semua atribut berperan penting dalam penelitian.

3. Data Cleaning (Pembersihan Data)

Proses pembersihan data dengan cara menghapus duplikasi data atau data yang berganda, memperbaiki data yang salah, dan mengecek data yang tidak konsisten.

4. Data Transformation

Pada penelitian ini, proses *transformation* dengan mengubah nilai data menjadi katagori, sehingga memudahkan dalam proses pembentukan pohon keputusan. Berikut data hasil proses *selection, cleaning, dan transformation*:

Tabel 3. Data Olah

Umur	anaemia	CPK	Diabetes	fraksi ejeksi	Hipertensi	trombosit	Tingkat Kreatinin	Tingkat Natrium	jkel	merokok	waktu tindakan	Harapan Hidup
------	---------	-----	----------	---------------	------------	-----------	-------------------	-----------------	------	---------	----------------	---------------

71-80	tidak	>192	Tidak	<55	Ya	150000-450000	<1.2	<135	p	tidak	0-20	Ya
51-60	tidak	>192	Tidak	<55	Tidak	150000-450000	<1.2	135-145	p	tidak	0-20	Ya
61-70	tidak	26-192	Tidak	<55	Tidak	150000-450000	<1.2	<135	p	ya	0-20	Ya
40-50	ya	26-192	Tidak	<55	Tidak	150000-450000	<1.2	135-145	p	tidak	0-20	Ya
61-70	ya	39-308	Ya	<55	Tidak	150000-450000	>1.4	<135	l	tidak	0-20	Ya
81-90	ya	26-192	Tidak	<55	Ya	150000-450000	>1.2	<135	p	ya	0-20	Ya
71-80	ya	>192	Tidak	<55	Tidak	<150000	<1.2	135-145	p	tidak	0-20	Ya
51-60	ya	>192	Ya	55-70	Tidak	>450000	<1.2	<135	p	ya	0-20	Ya
61-70	tidak	39-308	Tidak	55-70	Tidak	150000-450000	<1.4	135-145	l	tidak	0-20	Ya
71-80	ya	26-192	Tidak	<55	Ya	150000-450000	>1.2	<135	p	ya	0-20	Ya
71-80	ya	26-192	Tidak	<55	Ya	150000-450000	>1.2	<135	p	ya	0-20	Ya
61-70	tidak	>192	Tidak	<55	Ya	150000-450000	<1.2	135-145	p	ya	0-20	Ya
40-50	ya	>192	Tidak	<55	Tidak	<150000	<1.2	135-145	p	tidak	0-20	Ya
40-50	ya	26-192	Tidak	<55	Ya	150000-450000	<1.2	135-145	p	tidak	0-20	Ya
40-50	ya	39-308	Tidak	<55	Ya	150000-450000	<1.4	135-145	l	tidak	0-20	Tidak
81-90	ya	>192	Tidak	<55	Tidak	<150000	<1.2	135-145	p	tidak	0-20	Ya
81-90	ya	26-192	Tidak	<55	Tidak	150000-450000	<1.2	135-145	p	tidak	0-20	Ya

2.4 Pemanfaatan Algoritma Data Mining

Pada penelitian ini menggunakan algoritma yang ada pada data mining, yaitu algoritma C4.5 dan algoritma C4.5 dengan algoritma PSO (*Particle Swarm Optimization*). Algoritma C4.5 disebut juga dengan algoritma *decission tree* yang terkenal pada klasifikasi [8]. Kelebihan algoritma C4.5 [9] adalah dapat menghindari munculnya permasalahan yang terjadi pada pohon keputusan dengan memakai kriteria yang jumlahnya sedikit pada setiap node tanpa menguangi keputusan yang dihasilkan. Algoritma PSO (*Particle Swarm Optimization*) dikenal dengan sebutan algoritma optimasi yang dapat dipai untuk seleksi atribut. Dibanding dengan algoritma genetika lainnya, algoritma PSO bekerja lebih baik dalam bidang optimasi [8]. Kelebihan dari algoritma PSO yaitu mudah diterapkan, efisiensi saat melakukan proses pengolahan, dan memiliki konsep yang sederhana.

2.5 Evaluasi

Proses evaluasi pada penelitian ini menggunakan kurva *Receiver Operating Characteristic* (ROC) dan akurasi menggunakan *Confusion Matrix*. Terdapat lima kelompok klasifikasi tingkat ROC/AUC, yaitu:[10]

- a. *Excellent Classification*, memiliki nilai akurasi 0.90-1.00
- b. *Good Classification*, memiliki nilai akurasi 0.80-0.90
- c. *Fair Classification*, memiliki nilai akurasi 0.70-0.80
- d. *Poor Classification*, memiliki nilai akurasi 0.60-0.70
- e. *Failure*, memiliki nilai akurasi 0.50-0.60.

2.6 Perancangan Aplikasi Harapan Hidup Penyakit Gagal Jantung

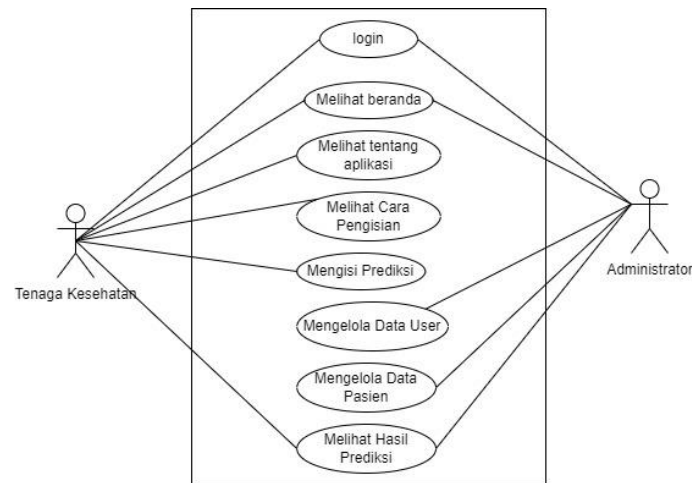
Perancangan aplikasi diawali dengan menganalisa kebutuhan user. Pada rancangan aplikasi ini terdapat dua user, yaitu tenaga kesehatan dan administrator. Berikut analisa kebutuhan dari tenaga kesehatan:

- a. Melihat beranda
- b. Melihat tentang aplikasi
- c. Melihat cara pengisian
- d. Mengisi data prediksi harapan hidup pasien gagal jantung dan melihat hasilnya

Berikut analisa kebutuhan dari administrator:

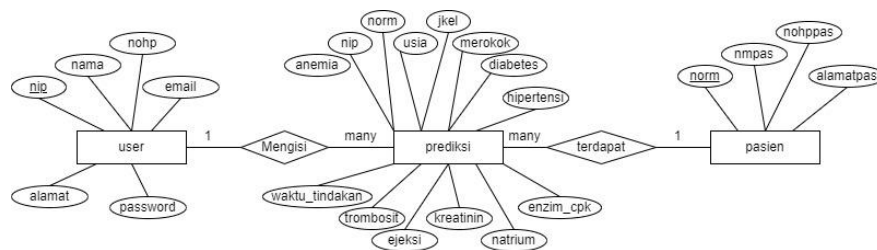
- a. Melihat beranda dan profil user
- b. Mengelola data user (tambah, edit, hapus)
- c. Mengelola data pasien (tambah dan edit)
- d. Melihat hasil prediksi harapan hidup pasien gagal jantung

Setelah mengetahui kebutuhan masing-masing user, selanjutnya merancang usecase diagram. Usecase diagram ini berfungsi untuk memberikan gambaran hubungan antar user dan sistem aplikasi yang dirancang. Berikut usecase diagramnya:



Gambar 2. Usecase Diagram

Setelah merancang usecase diagram, langkah selanjutnya merancang database yang dibutuhkan, seperti dibawah ini:



Gambar 3. Perancangan Database

3. HASIL DAN PEMBAHASAN

Pada penelitian ini mencari rule prediksi harapan hidup penyakit gagal jantung, metode yang dipakai algoritma C4.5 dan algoritma C4.5 dengan PSO (*Particle Swarm Optimization*). Setelah itu, melakukan proses pengujian berupa kurva *Receiver Operating Characteristic* (ROC) dan *Confusion Matrix*. Proses pengujian pertama menggunakan algoritma C4.5 dan pengujian kedua menggunakan algoritma C4.5 dengan PSO (*Particle Swarm Optimization*). Pencarian rule dan pengujian ini menggunakan *RapidMiner*.

3.1 Pencarian Rule Algoritma C4.5

Pencarian rule pertama menggunakan algoritma C4.5 dengan bantuan aplikasi RapidMiner. Hasil pemodelan diatas dengan menggunakan data sebesar 299 record. Berikut beberapa hasil rule pengolahan algoritma C4.5 dan PSO seperti dibawah:

- 1) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 40-50 tahun, maka tidak ada harapan hidup.
- 2) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 61-70 tahun, tidak mengalami anaemia, maka masih terdapat harapan hidup.
- 3) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 61-70 tahun, mengalami anaemia, kadar CPK antara 39-308, fraksi enjeksi <55, mempunyai hipertensi, trombosit antara 150000-450000, tingkat kreatinin <1.4, tingkat natrium antara 135-145, tidak merokok, maka masih terdapat harapan hidup.
- 4) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, mempunyai diabetes, maka tidak ada harapan hidup.
- 5) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu perempuan, maka masih terdapat harapan hidup.
- 6) Jika waktu tindakan 101-120 detik, umur antara 40-50, maka tidak ada harapan hidup.
- 7) Jika waktu tindakan 101-120 detik, umur antara 51-60, maka tidak ada harapan hidup.
- 8) Jika waktu tindakan 101-120 detik, umur antara 61-70, maka tidak ada harapan hidup.
- 9) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK antara 39-308, maka ada harapan hidup.
- 10) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK >192, tingkat natrium antara 135-145, maka tidak ada harapan hidup.
- 11) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK >192, tingkat natrium antara <135, maka masih terdapat harapan hidup.
- 12) Jika waktu tindakan 101-120 detik, umur antara 71-80, mengalami anaemia, maka masih terdapat harapan hidup.
- 13) Jika waktu tindakan 101-120 detik, umur antara 81-90, maka tidak ada harapan hidup.
- 14) Jika waktu tindakan 121-140 detik, tidak memiliki diabetes, maka tidak ada harapan hidup.
- 15) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 40-50, maka masih terdapat harapan hidup.
- 16) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 51-60, maka masih terdapat harapan hidup.
- 17) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 61-70, maka tidak ada harapan hidup.
- 18) Jika waktu tindakan 141-160 detik, tingkat kreatinin <1.2, tidak memiliki diabetes, maka tidak ada harapan hidup.
- 19) Jika waktu tindakan 141-160 detik, tingkat kreatinin <1.2, memiliki diabetes, tidak merokok, maka tidak ada harapan hidup.
- 20) Jika waktu tindakan 141-160 detik, tingkat kreatinin <1.2, memiliki diabetes, merokok, kadar CPK antara 26-192, maka masih terdapat harapan hidup.

3.2 Pencarian Rule Algoritma C4.5 dan PSO (Particle Swarm Optimization)

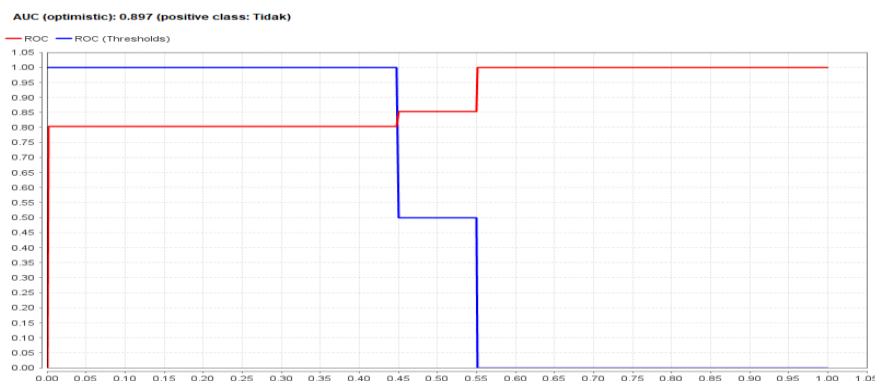
Pencarian rule selanjutnya menggunakan algoritma C4.5 dan PSO (*Particle Swarm Optimization*) dengan bantuan aplikasi RapidMiner. Hasil pemodelan diatas dengan menggunakan data sebesar 299 record. Berikut beberapa hasil rule pengolahan algoritma C4.5 dan PSO seperti dibawah:

- 1) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 40-50 tahun, maka tidak ada harapan hidup.
- 2) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 61-70 tahun, tidak mengalami anaemia, maka masih terdapat harapan hidup.

- 3) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, tidak mempunyai diabetes, umur antara 61-70 tahun, mengalami anaemia, kadar CPK antara 39-308, fraksi ejeksi <55, mempunyai hipertensi, trombosit antara 150000-450000, tingkat kreatinin <1.4, tingkat natrium antara 135-145, tidak merokok, maka masih terdapat harapan hidup.
- 4) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu laki-laki, mempunyai diabetes, maka tidak ada harapan hidup.
- 5) Jika waktu tindakan 0-20 detik, jenis kelamin yaitu perempuan, maka masih terdapat harapan hidup.
- 6) Jika waktu tindakan 101-120 detik, umur antara 40-50, maka tidak ada harapan hidup.
- 7) Jika waktu tindakan 101-120 detik, umur antara 51-60, maka tidak ada harapan hidup.
- 8) Jika waktu tindakan 101-120 detik, umur antara 61-70, maka tidak ada harapan hidup.
- 9) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK antara 39-308, maka ada harapan hidup.
- 10) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK >192, tingkat natrium antara 135-145, maka tidak ada harapan hidup.
- 11) Jika waktu tindakan 101-120 detik, umur antara 71-80, tidak mengalami anaemia, kadar CPK >192, tingkat natrium antara <135, maka masih terdapat harapan hidup.
- 12) Jika waktu tindakan 101-120 detik, umur antara 71-80, mengalami anaemia, maka masih terdapat harapan hidup.
- 13) Jika waktu tindakan 101-120 detik, umur antara 81-90, maka tidak ada harapan hidup.
- 14) Jika waktu tindakan 121-140 detik, tidak memiliki diabetes, maka tidak ada harapan hidup.
- 15) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 40-50, maka masih terdapat harapan hidup.
- 16) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 51-60, maka masih terdapat harapan hidup.
- 17) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 61-70, maka tidak ada harapan hidup.
- 18) Jika waktu tindakan 121-140 detik, memiliki diabetes, umur antara 71-80, maka masih terdapat harapan hidup.
- 19) Jika waktu tindakan 141-160 detik, tingkat kreatinin <1.2, tidak memiliki diabetes, maka tidak ada harapan hidup.
- 20) Jika waktu tindakan 141-160 detik, tingkat kreatinin <1.2, memiliki diabetes, tidak merokok, maka tidak ada harapan hidup.

3.3 Proses Performance Dengan Algoritma C4.5

Hasil *performance* kurva ROC pada algoritma C4.5, *false positive* dinyatakan garis horizontal dan *true positive* dinyatakan dengan garis vertical. Nilai dari grafik ROC yang didapat sebesar 0.897, sehingga termasuk dalam *good classification*. Karena nilai 0.897 terdapat pada rentang 0.80-0.90. Berikut Grafik dari ROC:



Gambar 3. Grafik ROC atau AUC Algoritma C4.5

Saat menggunakan algoritma C4.5 didapat nilai akurasi sebesar sebesar 73,33%. Berdasarkan 299 data didapatkan 16 data prediksi masih ada harapan hidup sesuai real dengan ya (masih terdapat harapan hidup), 11 prediksi masih ada harapan hidup ternyata tidak (tidak terdapat harapan hidup), 13 data prediksi tidak ada harapan hidup ternyata masih ada harapan hidup, dan 50 data prediksi tidak ada harapan hidup sesuai dengan tidak ada harapan hidup. Dengan persentase prediksi yang masih terdapat harapan hidup sebesar 59.26% dan persentase prediksi yang tidak ada harapan hidup sebesar 79.37%. Persentase *class recall* pada yang masih ada harapan hidup sebesar 55.17% dan persentase *class recall* yang benar tidak ada harapan hidup sebesar 81.97%. Berikut gambar tabel persentase akurasi:

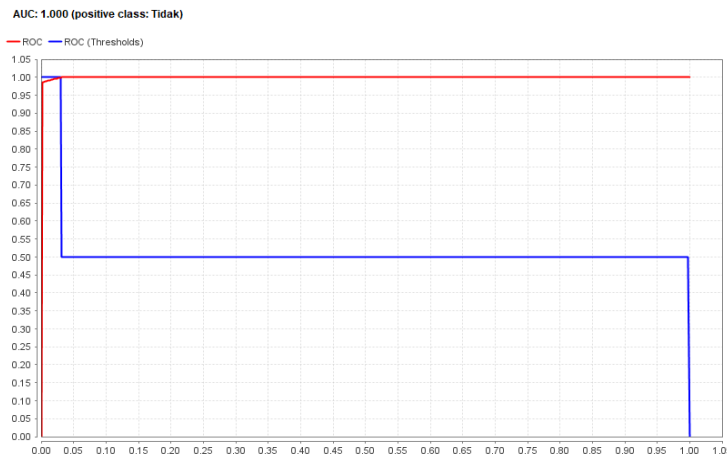
accuracy: 73.33%

	true Ya	true Tidak	class precision
pred. Ya	16	11	59.26%
pred. Tidak	13	50	79.37%
class recall	55.17%	81.97%	

Gambar 4. Tingkat Akursi *Confusion Matrix* Pada Algoritma C4.5

3.4 Proses Performance Dengan Algoritma C4.5 dan PSO (*Particle Swarm Optimization*)

Hasil *performance* kurva ROC pada algoritma C4.5 dan PSO (*Particle Swarm Optimization*), *false positive* dinyatakan garis horizontal dan *true positive* dinyatakan dengan garis vertical. Nilai dari grafik ROC yang didapat sebesar 1.000, sehingga termasuk dalam *Excellent Classification*. Karena nilai 1.000 terdapat pada rentang 0.90-1.00. Berikut Grafik dari ROC:



Gambar 5. Grafik ROC atau AUC Algoritma C4.5 dan PSO

Saat menggunakan algoritma C4.5 dan PSO (*Particle Swarm Optimization*) didapat nilai akurasi sebesar sebesar 99.00%. Berdasarkan 299 data didapatkan 95 data prediksi masih ada harapan hidup sesuai real dengan ya (masih terdapat harapan hidup), 2 prediksi masih ada harapan hidup ternyata tidak (tidak terdapat harapan hidup), 1 data prediksi tidak ada harapan hidup ternyata masih ada harapan hidup, dan 201 data prediksi tidak ada harapan hidup sesuai dengan tidak ada harapan hidup. Dengan persentase prediksi yang masih terdapat harapan hidup sebesar 97.94% dan persentase prediksi yang tidak ada harapan hidup sebesar 99.50%. Persentase *class recall* pada yang masih ada harapan hidup sebesar 98.96% dan persentase *class recall* yang benar tidak ada harapan hidup sebesar 99.01%. Berikut gambar tabel persentase akurasi:

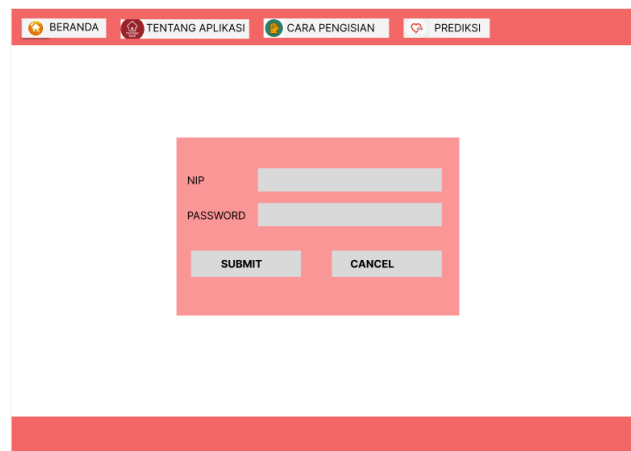
accuracy: 99.00%

	true Ya	true Tidak	class precision
pred. Ya	95	2	97.94%
pred. Tidak	1	201	99.50%
class recall	98.96%	99.01%	

Gambar 6. Tingkat Akursi *Confussion Matrix* Pada Algoritma C4.5

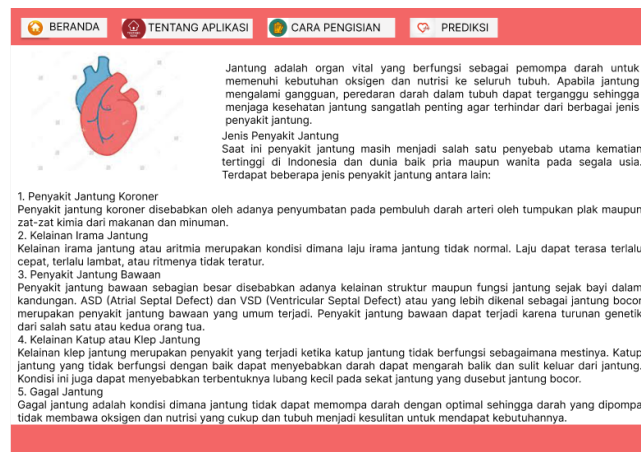
3.5 Perancangan Aplikasi Prediksi Harapan Hidup Pasien Gagal Jantung

Rancangan aplikasi prediksi ini terdapat dua user, yaitu Administrator dan Tenaga Kesehatan. Administrator yang berfungsi mengelola jalannya aplikasi dan Tenaga Kesehatan melakukan input prediksi. Berikut tampilan dari halaman login Tenaga Kesehatan dan Administrator:



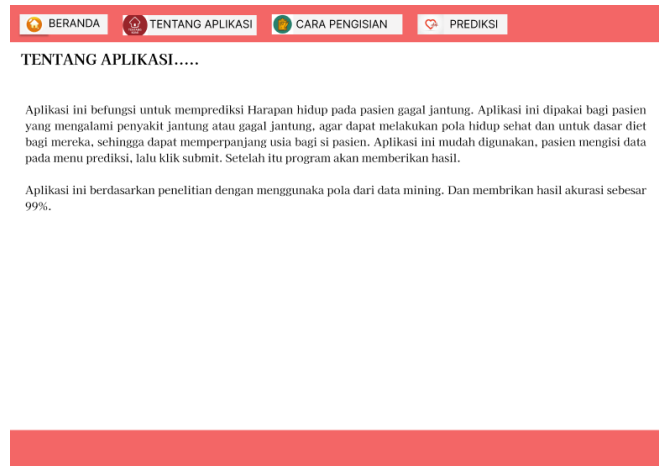
Gambar 7. Halaman Login

Setelah user tenaga kesehatan melakukan login, maka akan menuju ke beranda. Berikut tampilan beranda:



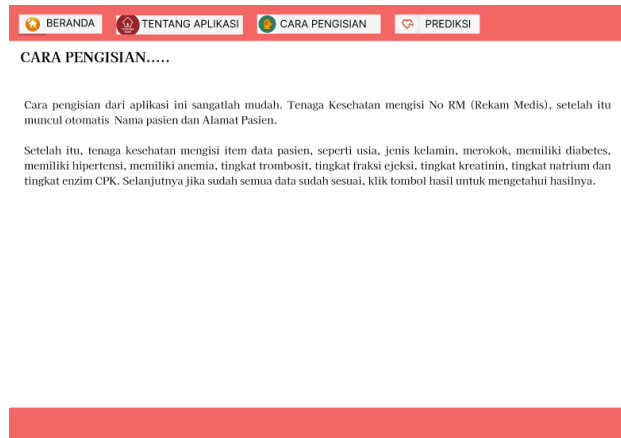
Gambar 8. Halaman Beranda Tenaga Kesehatan

Halaman beranda ini berisi artikel kesehatan jantung. Selain beranda, terdapat halaman lainnya yaitu tentang aplikasi, cara pengisian dan prediksi. Berikut tampilan halaman tentang aplikasi pada user tenaga kesehatan:



Gambar 9. Halaman Tentang Aplikasi Pada Tenaga Kesehatan

Selanjutnya tampilan halaman cara pengisian aplikasi pada user tenaga kesehatan, sebagai berikut:



Gambar 10. Halaman Cara Pengisian Aplikasi Pada Tenaga Kesehatan

Selanjutnya tampilan halaman prediksi aplikasi pada user tenaga kesehatan. Pada halaman prediksi ini, tenaga kesehatan dapat mengisi kondisi pasien yang sedang mengalami gagal jantung saat melakukan pengobatan. Setelah mengisi, lalu klik hasil untuk menampilkan hasil dari prediksi. Berikut tampilan halaman prediksi seperti di bawah ini:



Gambar 11. Halaman Prediksi Pada Tenaga Kesehatan

Selanjutnya tampilan halaman hasil prediksi pada user tenaga kesehatan, sebagai berikut:

The screenshot displays a prediction results page for a healthcare user. At the top, there is a navigation bar with four items: BERANDA, TENTANG APLIKASI, CARA PENGISIAN, and PREDIKSI. The main content area is divided into two columns. The left column contains patient information: Nomer RM (001/OM/20211005-001), Nama Pasien (Anwar Rohim), Alamat Pasien (Jl. Cemerlang No. 99, RT.11 RW.02, Cipayung), Usia Pasien (67), Jenis Kelamin (Laki-Laki), Merokok (Ya), Diabetes (Ya), Hipertensi (Tidak), and Anemia (Tidak). The right column contains lab test results: Periode Waktu Tindakan (121-140), Tingkat Fraksi Ejeksi (55-70), Tingkat Kreatinin (>1.2), Tingkat Natrium (>145), Tingkat Enzim CPK (26-192), and Tingkat Trombosit (150000-450000). At the bottom right, there is a 'HASIL' section with a warning: 'Untuk harapan hidup sangat kecil. Oleh karena itu pasien harus menerapkan hidup sehat dan berolah raga'.

Gambar 12. Halaman Prediksi Pada Tenaga Kesehatan

Pada halaman login, dapat digunakan oleh 2 user, yaitu administrator dan tenaga kesehatan. Pada user administrator, setelah melakukan login akan menuju ke beranda administrator. Berikut halaman beranda dari user administrator:

The screenshot shows the administrator's home page. At the top right, the user's name and ID are displayed: 'Qeela Arsyad, PSpJ220305012'. Below this, there is a navigation menu with four items: BERANDA, DATA USER, DATA PASIEN, and LIHAT PREDIKSI. The main content area is titled 'Selamat Datang, Qeela Arsyad' and 'Profile'. The profile section contains the following details: NIP (PSpJ220305012), NAMA (Qeela Arsyad), EMAIL (Qeela.43@gmail.com), NO. HP (0852874579037), and ALAMAT (Jl. Mawar Raya No.14, RT.07 RW.02, Cijantung).

Gambar 13. Halaman Beranda Pada Administrator

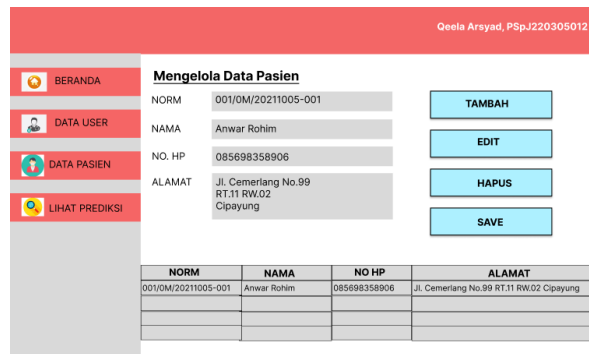
Berikut halaman mengelola data user pada user administrator. Pada halaman ini, administrator melakukan isi data user, seperti nip, nama, email, no hp alamat dan password. Pada halaman ini bisa menambahkan data, mengedit dan menghapus data.

The screenshot displays the 'Mengelola Data User' page. At the top right, the user's name and ID are shown: 'Qeela Arsyad, PSpJ220305012'. The navigation menu includes BERANDA, DATA USER, DATA PASIEN, and LIHAT PREDIKSI. The main content area is titled 'Mengelola Data User' and contains a form with the following fields: NIP (PSpJ220305012), NAMA (Qeela Arsyad), EMAIL (Qeela.43@gmail.com), NO. HP (0852874579037), ALAMAT (Jl. Mawar Raya No.14, RT.07 RW.02, Cijantung), and PASSWORD (123456). To the right of the form are four buttons: TAMBAH, EDIT, HAPUS, and SAVE. Below the form is a table with the following data:

NIP	NAMA	EMAIL	NO.HP	NO.HP
PSpJ220305012	Qeela Arsyad	Qeela.43@gmail	0852874579037	Jl. Mawar Raya

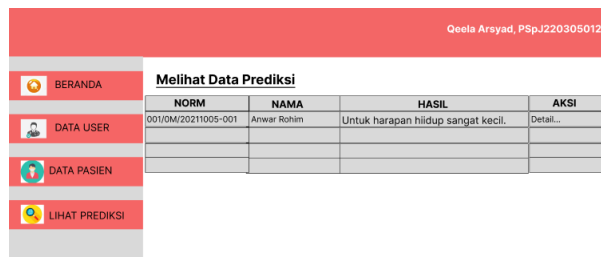
Gambar 14. Halaman Mengelola Data User Pada Administrator

Berikut halaman mengelola data pasien pada user administrator. Pada halaman ini, administrator melakukan isi data pasien, seperti nomer rekam medik, nama pasien, no hp, dan alamat pasien. Pada halaman ini bisa menambahkan data, mengedit dan menghapus data.



Gambar 15. Halaman Mengelola Data Pasien Pada Administrator

Berikut halaman lihat hasil prediksi pada administrator. Pada halaman ini, administrator hanya bias melihat hasil prediksi saja:



Gambar 16. Halaman Melihat Hasil Prediksi Pada Administrator

4. KESIMPULAN

Kesimpulan dari pembahasan perbandingan algoritma c4.5 dengan algoritma c4.5 dan pso sebagai prediksi harapan hidup pasien gagal jantung, yaitu:

- Kriteria lama waktu tindakan memiliki prioritas utama dalam prediksi penyakit gagal jantung.
- Dari hasil pengolahan data, dengan menggunakan algoritma C4.5 didapatkan akurasi sebesar 73.33%, sedangkan menggunakan algoritma C4.5 dan PSO didapatkan akurasi sebesar 99.00%.
- Sehingga dapat disimpulkan berdasarkan tingkat akurasi bahwa pemodelan algoritma C4.5 dan PSO memiliki keakuratan yang lebih tinggi dibandingkan algoritma C4.5.
- Dari hasil pengolahan data, dengan menggunakan algoritma C4.5 didapatkan akurasi grafik ROC sebesar 0.897%, sedangkan menggunakan algoritma C4.5 dan PSO didapatkan akurasi grafik ROC sebesar 1.00%.
- Sehingga dapat disimpulkan berdasarkan tingkat akurasi grafik ROC bahwa pemodelan algoritma C4.5 dan PSO memiliki keakuratan yang lebih tinggi dibandingkan algoritma C4.5.
- Dengan didapatnya hasil akurasi tertinggi yaitu algoritma C4.5 dan PSO. Oleh karena itu dibuat rancangan aplikasinya, agar memberikan kemudahan dalam mendapatkan hasil prediksi.

Untuk pengembangan selanjutnya, penelitian ini dapat dibuatkan aplikasinya agar menjadi perubahan gaya hidup bagi yang mengalami penyakit gagal jantung.

REFERENSI

- [1] Rokom, "Penyakit Jantung Koroner Didominasi Masyarakat Kota," *kemkes.go.id*, 2021. <https://sehatnegeriku.kemkes.go.id/>
- [2] D. P. T. Astuti and I. K. Suardamana, "Gagal Jantung Tinjauan pustaka," *Ilmu Penyakit Dalam*, no. 1002005139, p. 1513, 2017.
- [3] M. Beyer, R. Lenz, and K. A. Kuhn, *Health Information Systems*, vol. 48, no. 1. 2006. doi: 10.1524/itit.2006.48.1.6.
- [4] 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.
- [5] F. Novaldy and A. Herliana, "Penerapan Pso Pada Naïve Bayes Untuk Prediksi Harapan Hidup Pasien Gagal Jantung," *J. Responsif Ris. Sains dan Inform.*, vol. 3, no. 1, pp. 37–43, 2021, doi: 10.51977/jti.v3i1.396.
- [6] M. Yunus, H. Ramadhan, D. R. Aji, and A. Yulianto, "Penerapan Metode Data Mining C4.5 Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP)," *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 2, 2021, doi: 10.31294/p.v23i2.11395.
- [7] R. S. Asa, S. Defit, and J. Na'am, "Identifikasi Penyaluran Zakat Menggunakan Algoritma C4.5 (Studi Kasus Di Baznas Kabupaten Agam)," *J. Sains dan Inform.*, vol. 4, no. 1, pp. 44–53, 2019, doi: 10.22216/jsi.v4.
- [8] D. Safira and Mustakim, "Perbandingan Algoritma C4.5 dengan C4.5+Particle Swarm Optimization untuk Klasifikasi Angkatan Kerja," *J. Politek. Caltex Riau*, vol. 7, no. 2, pp. 272–279, 2021.
- [9] S. Ucha Putri, E. Irawan, F. Rizky, S. Tunas Bangsa, P. A. -Indonesia Jln Sudirman Blok No, and S. Utara, "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5," *Januari*, vol. 2, no. 1, pp. 39–46, 2021.
- [10] E. Fitriani, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan," *Sistemasi*, vol. 9, no. 1, p. 103, 2020, doi: 10.32520/stmsi.v9i1.596.