

KOMPARASI ALGORITMA C4.5 DAN NAIVE BAYES DALAM PREDIKSI PENDERITA PENYAKIT GAGAL JANTUNG

Robby Anggriawan^{1,*}, Handoyo Widi Nugroho²

ABSTRACT

Every year more than 36 million people die from Non-Communicable Diseases (PTM) (63% of all deaths). More than 9 million deaths from non-communicable diseases occur before the age of 60, and 90% of these “early” deaths occur in low- and middle-income countries. Globally PTM the number one cause of death each year is cardiovascular disease. Cardiovascular disease is a disease caused by impaired function of the heart and blood vessels, such as: coronary disease, heart failure or heart failure, hypertension and stroke. An estimated 17.3 million deaths are caused by cardiovascular disease.

Keywords: C45 Algorithm Naïve Bayes algorithm; Rapid Miner;

ABSTRAK

Setiap tahunnya lebih dari 36 juta orang meninggal karena Penyakit Tidak Menular (PTM) (63% dari seluruh kematian). Lebih dari 9 juta kematian yang disebabkan oleh penyakit tidak menular terjadi sebelum usia 60 tahun, dan 90% dari kematian “dini” tersebut terjadi di negara berpenghasilan rendah dan menengah. Secara Global PTM penyebab kematian nomor satu setiap tahunnya adalah penyakit kardiovaskuler. Penyakit Kardiovaskuler adalah penyakit yang disebabkan gangguan fungsi jantung dan pembuluh darah, seperti: penyakit koroner, penyakit gagal jantung atau payuh jantung, hipertensi dan stroke. Diperkirakan sebanyak 17,3 juta kematian disebabkan oleh penyakit kardiovaskuler.

Kata Kunci: Algoritma C45; Algoritma Naïve Bayes; Rapid Miner

1. PENDAHULUAN

Salah satu penyakit yang sering terjadi dan tidak menular (PTM) adalah penyakit jantung. Penyakit jantung tersebut merupakan penyakit merupakan kondisi yang terjadi ketika pembuluh darah utama yang menyuplai darah ke jantung mengalami kerusakan dan tidak dapat bekerja dengan semestinya hal tersebut dikarenakan banyak faktor. Tumpukan kolesterol pada pembuluh darah serta proses peradangan diduga salah satu menjadi faktor penyebab penyakit ini. Ketika terjadi penumpukan kolesterol (plak), pembuluh darah koroner akan menyempit sehingga aliran darah dan suplai oksigen menuju jantung pun akan terhambat sehingga mempengaruhi kestabilan proses bekerja jantung. Kurangnya aliran darah ini akan menyebabkan rasa nyeri pada dada (angina) dan sesak napas, hingga suatu saat terjadi hambatan total pada aliran darah menuju jantung atau yang disebut juga dengan serangan jantung. Menurut banyak penelitian penyakit jantung koroner termasuk salah satu penyebab kematian tertinggi di dunia.

Federasi Jantung Dunia memperkirakan angka kematian akibat penyakit jantung koroner di Asia Tenggara mencapai 1,8 juta kasus pada tahun 2014. Di Indonesia sendiri pada tahun 2013 tercatat ada setidaknya 883.447 orang yang terdiagnosis penyakit jantung koroner di Indonesia dengan mayoritas penderita berusia 55-64 tahun. Angka

kematian akibat penyakit jantung pun menjadi cukup tinggi, yakni sekitar 45 persen dari seluruh angka kematian di Indonesia (Subarkah, Risma, and Aditya 2022)

Pada acara peringatan hari jantung sedunia (*world heart day*) yang dicetuskan pertama kali oleh World Heart Federation pada tahun 2000 menginformasikan kepada orang-orang di seluruh dunia bahwa penyakit jantung dan stroke merupakan penyebab utama kematian di dunia yang saat ini di klaim mencapai 17,3 juta kematian setiap tahunnya. Angka kematian akibat penyakit jantung tersebut diperkirakan akan terus meningkat hingga mencapai 23,3 juta pada tahun 2030. Berdasarkan Hasil Riset Kesehatan Dasar (Riskesdas) Kemenkes RI Tahun 2013, prevalensi penyakit jantung koroner di Indonesia mencapai 0,5% dan gagal jantung sebesar 0,13% dari total penduduk berusia 18 tahun keatas sehingga harus penyakit tersebut harus selalu diwaspadai (Informatika et al. 2022) Banyak faktor yang dapat meningkatkan risiko terkena penyakit jantung. Faktor risiko tersebut terdiri dari faktor risiko yang tidak dapat dimodifikasi seperti riwayat keluarga, umur serta jenis kelamin dan faktor risiko yang dapat dimodifikasi seperti hipertensi, kebiasaan merokok, penyakit diabetes, dislipidemia, obesitas, kurang aktifitas fisik, pola makan serta stres. Penyakit pembuluh darah, atau disebut dengan penyakit kardiovaskular atau disebut juga penyakit jantung. Penyakit ini berhubungan dengan proses aterosklerosis, yaitu suatu kondisi pada organ tubuh yang terjadi ketika zat yang disebut plak menumpuk di dinding arteri. Penumpukan ini menyebabkan penyempitan pada arteri sehingga aliran darah terhenti atau tersumbat dan tidak dapat beredar dengan semestinya, hal ini dapat mengakibatkan serangan jantung atau stroke (Nurlia and Enri 2021)

Salah satu bagian penting dari pengobatan atau tindakan medis adalah pengambilan keputusan dan proses klasifikasi atau prediksi pada suatu hal yang menjadi fokus seperti pendeteksian penyakit, namun klasifikasi medis atau prediksi biasanya merupakan proses yang sangat kompleks dan sulit dilakukan jika tak dapat mengetahui metode yang tepat dan terbaik dalam memberikan solusinya (Hasanah, Oktavianto, and Rahayu 2022) Tantangan yang dihadapi oleh organisasi kesehatan adalah mendiagnosa pasien dengan benar, diagnosa atau prediksi yang buruk dapat menyebabkan konsekuensi yang mendatangkan malapetaka yang kemudian tidak dapat diterima (Putra and Rini 2019). Untuk menjawab tantangan tersebut beberapa penelitian telah dilakukan dalam bidang kesehatan untuk mendapatkan prediksi penyakit dengan lebih akurat, namun belum diketahui metode apa yang paling akurat dalam memprediksi penyakit pasien. Berbagai macam teknik analisa yang secara konvensional dan manual yang selama ini digunakan tidak lagi begitu efektif digunakan untuk hal mendiagnosa suatu penyakit. Seiring dengan perkembangan teknologi dan system. berbasis pengetahuan terutama medis tuntutan akan adanya penggunaan sistem pengetahuan berbasis komputer sebagai teknik analisa dalam mendiagnosa penyakit menjadi semakin penting dan harus selalu dikembangkan. Oleh karenanya, saat inilah waktu yang tepat untuk mengembangkan sistem pengetahuan berbasis komputer yang modern, efektif dan efisien dalam mendiagnosa masalah penyakit (Derisma 2020).

Oleh karena itu, penelitian ini dilakukan untuk membantu menyelesaikan permasalahan tersebut dengan data mining klasifikasi untuk prediksi penyakit jantung, diperlukan suatu metode atau teknik yang dapat mengolah data-data yang sudah ada. Salah satu metodenya menggunakan teknik data mining klasifikasi. Penggunaan data mining Algoritma C45 dan Naive Bayes dalam Klasifikasi Penderita Gagal Jantung sebagai pilihan untuk diagnosa penyakit jantung dapat menjadi alternatif pilihan yang tepat, tetapi sampai saat ini belum diketahui algoritma klasifikasi yang paling akurat dalam prediksi penyakit jantung.

Pada penelitian sebelumnya membandingkan algoritma klasifikasi data mining Naive Bayes Berbasis PSO untuk deteksi penyakit jantung. Pengukuran dengan Naives Bayes menghasilkan akurasi 82.14%, sementara dengan Naive Bayes Berbasis Particle Swarm Optimization akurasi meningkat menjadi 92.86%. (Putri and Huda 2020)

Pada penelitian algoritma menggabungkan k-NN dengan algoritma genetika untuk klasifikasi yang lebih efektif. Algoritma genetika melakukan proses yang kompleks dan berikan solusi optimal. Hasil percobaan menunjukkan bahwa algoritma kami Meningkatkan akurasi diagnosis penyakit jantung (Munandar and Munir 2022).

Berdasarkan penelitian tersebut penulis melakukan penelitian dengan melakukan komparasi data mining dengan metode Komparasi Algoritma C4.5 dan Naive Bayes dalam Klasifikasi Penderita Penyakit Gagal Jantung untuk mengetahui algoritma yang memiliki akurasi yang lebih tinggi dalam hal ini yaitu prediksi penyakit jantung.

2. KERANGKA TEORI

Penyakit jantung merupakan penyakit yang sangat berbahaya karena salah satu penyakit yang mematikan, terutama penyakit jantung koroner yang jumlah penderitanya sangat banyak dibandingkan dengan jenis penyakit jantung lainnya. Penyakit jantung koroner terjadi akibat adanya penyumbatan pembuluh arteri oleh plak yang menghambat suplai oksigen dan nutrisi ke jantung. Plak yang menyumbat pembuluh darah tidak muncul begitu saja. Kemunculan plak yang berupa timbunan lemak atau kalsium melalui proses secara bertahap. Biasanya, diawali dengan kekakuan pembuluh darah atau biasa di sebut dengan aterosklerosis, kemudian penyempitan pembuluh darah (Effendi and Muhamad 2021).

2.1. Gejala Penyakit Jantung

Penyakit jantung koroner atau arteri koroner dianggap sangat berbahaya karena dapat menimbulkan serangan jantung mendadak yang berujung kematian. Serangan jantung terjadi akibat terhambatnya aliran darah menuju jantung sehingga suplai oksigen dan nutrisi di otot jantung dan jaringan disekitar jantung berkurang.

Tidak seperti otot tubuh lainnya, otot jantung tidak memiliki kemampuan beregenerasi. Apabila terdapat saja kerusakan maka akan berakibat fatal bagi tubuh. Semakin lama serangan jantung terjadi semakin banyak pula kerusakan pada jantung. Karena itu penting bagi kita untuk mengenali gejala-gejala dari penyakit jantung sehingga dapat memberikan pertolongan dengan segera. Gejala pnuakit jantung koroner secaa umum tidak dikenali oleh orang awam. Mereka terkadang menyepelekan dan menganggapnya wajar. Penderita baru menyadari bahwa dirinya terkena penyakit jantung koroner ketika kondisinya sudah parah. Bahkan, tak jarang dari mereka pada akhirnya harus meregang nyawa karena keterlambatan penanganan (Nawawi, Purnama, and Hikmah 2019)

2.2. Data Mining

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan-perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada prediksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu (Dharmawan 2021). Kaka's data mining meramalkan tren dan sifat-sifat perilaku bisnis yang sangat berguna untuk mendukung pengambilan keputusan penting. Analisis yang diotomatisasi yang dilakukan oleh data mining melebihi yang dilakukan oleh sistem pendukung keputusan tradisional yang sudah banyak digunakan. Secara khusus, koleksi metode yang dikenal sebagai 'data mining' menawarkan metodologi dan solusi teknis untuk mengatasi analisis data medis dan konstruksi prediksi model (Informatika et al. 2022). Berdasarkan tugas dan tujuan analisis, proses data mining dapat dibagi menjadi dua kategori utama, Tergantung pada adanya target variabel dan metode belajar (*learning*) yaitu antara proses belajar yang diawasi (*supervised*) dan tanpa pengawasan (*unsupervised*) (Destiani et al. 2018).

2.3. Algoritma C4.5

Algoritma C4.5 adalah sebuah algoritma yang berfungsi untuk membangun *decision tree* (pohon keputusan). Algoritma C4.5 dan pohon keputusan merupakan dua model yang tidak terpisahkan. Algoritma C4.5 adalah salah satu dari algoritma klasifikasi yang kuat dan banyak digunakan atau di implementasikan untuk pengklasifikasian dalam berbagai hal. Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan (1996) sebagai versi perbaikan dari algoritma Iterative Dichotomiser 3 (ID3). Serangkaian perbaikan dilakukan pada algoritma ID3 mencapai puncaknya dengan menghasilkan sebuah sistem praktis dan simple yang berpengaruh untuk pembentukan pohon keputusan. Perbaikan tersebut meliputi metode untuk menangani data kontinyu, mengatasi missing data, dan melakukan pemangkasan pohon (Siahaan, Nasution, and Hasibuan 2021). Untuk menghitung nilai gain rumus yang digunakan adalah (hidayah nur umi and dkk 2021).

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^S \frac{S_i}{S} * Entropy(S_i)$$

2.4 Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan metode yang dapat digunakan untuk mengklasifikasikan sekumpulan data. Algoritma ini memanfaatkan metode probabilitas dan Statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya (Cahya Putri Buani 2021).

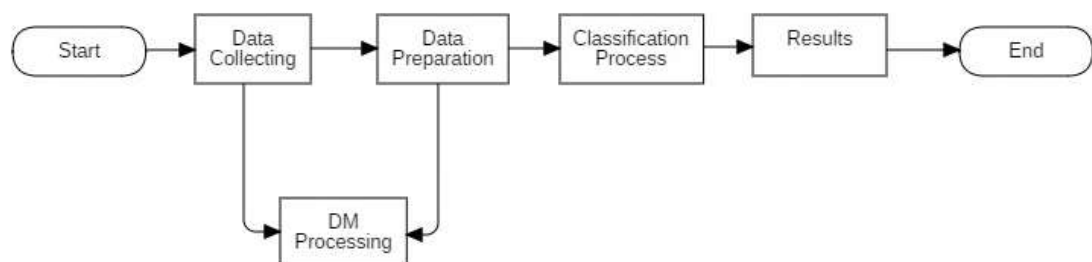
Naïve Bayes merupakan *machine learning* yang menggunakan perhitungan probabilitas yang menggunakan konsep pendekatan *Bayesian*. Kata *Naïve*, yang terkesan merendahkan, berasal dari asumsi *independensi* pengaruh nilai suatu atribut dari probabilitas pada kelas yang diberikan terhadap nilai atribut lainnya. Penggunaan teorema Bayes pada algoritma *Naïve Bayes* yaitu dengan mengkombinasikan prior probability dan probabilitas bersyarat dalam sebuah rumus yang bisa digunakan untuk menghitung probabilitas tiap klasifikasi yang mungkin.

Rumus *Naïve Bayes* nya adalah:

$$P(H|X) = \frac{P(H)P(X|H)}{P(X)}$$

3. METODOLOGI

Adapun tahapan penelitian yang digunakan merujuk pada gambar 1, sebagai berikut:



Gambar 1. Tahapan Penelitian

4. HASIL DAN PEMBAHASAN

Dalam bagian penulis akan membahas mengenai hasil dari penelitian, data tersebut akan dihitung menggunakan *feature selection* dengan menggunakan algoritma *Decision Tree C4.5* dan Algoritma Naive Bayes yang kemudian akan diuji menggunakan *Split Validation* menggunakan algoritma *Decision Tree C.45*. Algoritma Naive Bayes *Dataset* yang digunakan merupakan *dataset* publik yang berasal dari <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> dataset *public heart disease* yang di peroleh dari *dataset* dapat dilihat pada gambar berikut:

Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
40	M	ATA	140	289	0	Normal	172	N	0	Up	0
49	F	NAP	160	180	0	Normal	156	N	1	Flat	1
37	M	ATA	130	283	0	ST	98	N	0	Up	0
48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
54	M	NAP	150	195	0	Normal	122	N	0	Up	0
39	M	NAP	120	339	0	Normal	170	N	0	Up	0
45	F	ATA	130	237	0	Normal	170	N	0	Up	0
54	M	ATA	110	208	0	Normal	142	N	0	Up	0
37	M	ASY	140	207	0	Normal	130	Y	1.5	Flat	1
48	F	ATA	120	284	0	Normal	120	N	0	Up	0
37	F	NAP	130	211	0	Normal	142	N	0	Up	0
58	M	ATA	136	164	0	ST	99	Y	2	Flat	1
39	M	ATA	120	204	0	Normal	145	N	0	Up	0
49	M	ASY	140	234	0	Normal	140	Y	1	Flat	1
42	F	NAP	115	211	0	ST	137	N	0	Up	0
54	F	ATA	120	273	0	Normal	150	N	1.5	Flat	0
38	M	ASY	110	196	0	Normal	166	N	0	Flat	1
43	F	ATA	120	201	0	Normal	165	N	0	Up	0
60	M	ASY	100	248	0	Normal	125	N	1	Flat	1

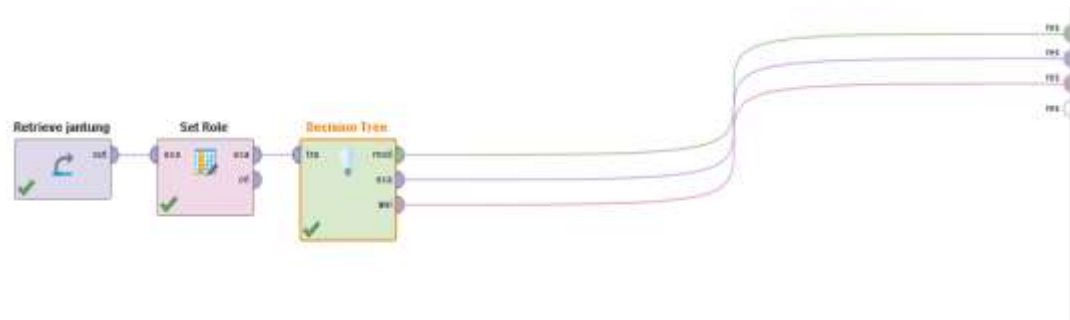
Gambar 2. Dataset Gagal Jantung

Data akan diseleksi untuk menentukan variabel apa saja yang akan diambil agar tidak terjadi kesamaan dan perulangan yang tidak diperlukan dalam pengolahan teknik data mining, dalam penanganan kasus yang terjadi pada *dataset public*, dalam penseleksian data dapat dilakukan dengan cara mengimport data ke dalam program rapidminer 9.10, data tersebut tidak akan dapat diolah karena masih terdapat beberapa *error* data dan program akan secara otomatis memberi tahu bahwa *dataset* perlu dilakukan *preprocessing*. Setelah data *error* tersebut ditemukan kita dapat menyeleksi dengan melakukan filter untuk menghapus data tersebut sehingga data dapat diproses kedalam sistem

Row No.	HeartDisease	Age	Sex	ChestPainTy...	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR
1	0	40	M	ATA	140	289	0	Normal	172
2	1	49	F	NAP	160	180	0	Normal	156
3	0	37	M	ATA	130	283	0	ST	98
4	1	48	F	ASY	138	214	0	Normal	108
5	0	54	M	NAP	150	195	0	Normal	122
6	0	39	M	NAP	120	339	0	Normal	170
7	0	45	F	ATA	130	237	0	Normal	170
8	0	54	M	ATA	110	208	0	Normal	142
9	1	37	M	ASY	140	207	0	Normal	130
10	0	48	F	ATA	120	284	0	Normal	120
11	0	37	F	NAP	130	211	0	Normal	142
12	1	58	M	ATA	136	164	0	ST	99
13	0	39	M	ATA	120	204	0	Normal	145
14	1	49	M	ASY	140	234	0	Normal	140

Gambar 3. Dataset Gagal Jantung

Pada algoritma *decision tree*, record yang sudah di *import* ke rapid miner akan digunakan untuk menentukan pola pohon keputusan, hasil dari pola pohon keputusan dapat dilihat pada gambar berikut.

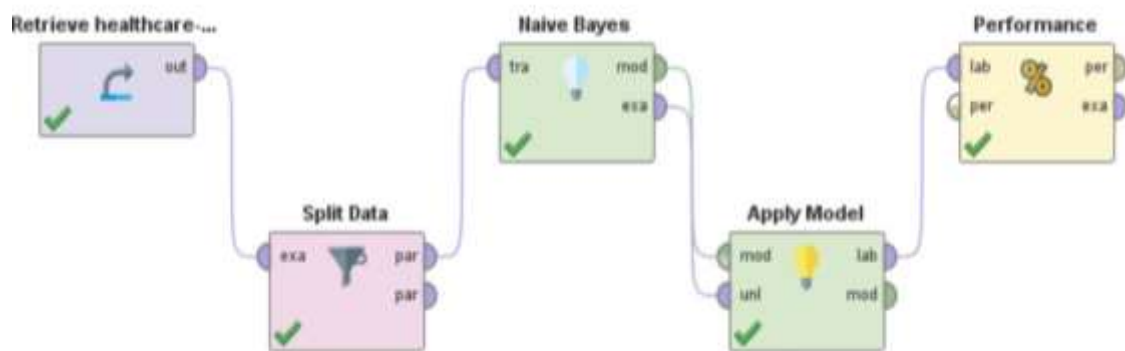


Gambar 4. Proses



Gambar 5. Pola Pohon Keputusan

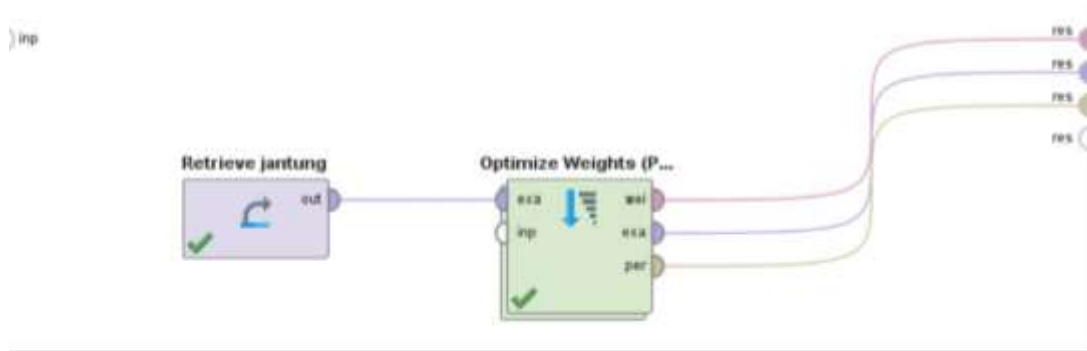
Penerapan data pada Rapidminer untuk klasifikasi dengan menggunakan *split validation* dengan nilai akurasi, *precision*, *confusion matrix* atau nilai *recall* dan nilai AUC dengan pembagian *data training* dan *testing* sebesar 70:30. dapat dilihat pada gambar 6.



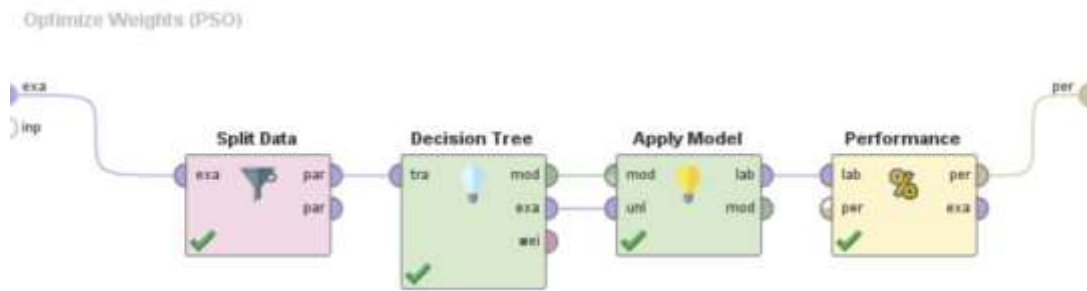
Gambar 6. Proses Menggunakan Algoritma Naïve Bayes *Split Validation*

4.1 Menggunakan *Particle Swarm Optimization (PSO)* *Decision Tree C4.5* dan Algoritma Naïve Bayes

Metode untuk meningkatkan tingkat akurasi dalam klasifikasi adalah penggunaan teknik optimasi. Sehingga peneliti menggunakan eksperimen lanjutan dengan menggunakan teknik optimasi seleksi fitur *Particle Swarm Optimization (PSO)* dengan algoritma *Decision Tree C4.5*. dan Algoritma Naïve Bayes Penerapan Metode dan algoritma pada rapidminer 9.10 ditunjukkan pada gambar 7.

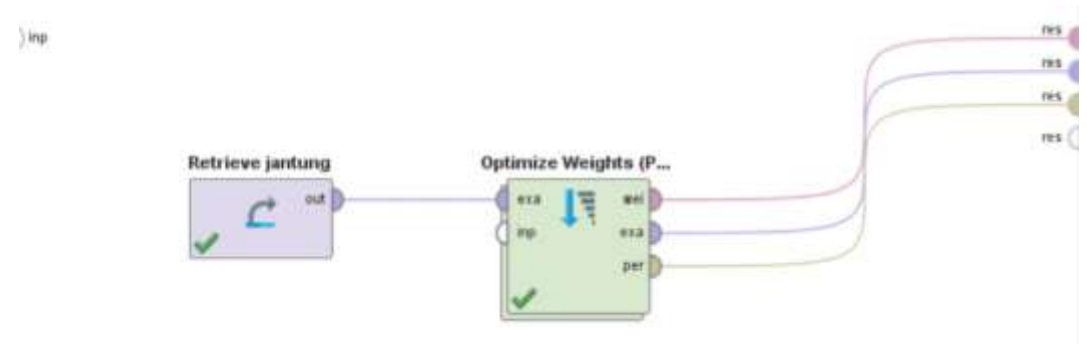


Gambar 7. Proses 1 Particle Swarm Optimization (PSO) dan Decision Tree C4.5

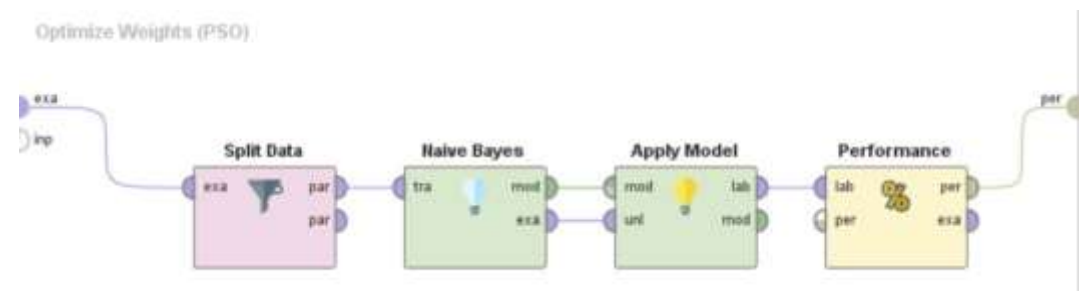


Gambar 8. Proses 2 Particle Swarm Optimization (PSO) dan Decision Tree C4.5

4.2 Particle Swarm Optimization (PSO) Naïve Bayes Dengan Split Data (70:30)



Gambar 9. Proses 1 Particle Swarm Optimization (PSO) dan Naïve Bayes



Gambar 10. Proses 1 Particle Swarm Optimization (PSO) dan Naïve Bayes

Tabel 1. Perbandingan Akurasi

No	Algoritma	Akurasi	AUC
1	C4.5	95,16%	0,773
2	C4.5 + PSO	95,18%	0,988
3	Naïve Bayes	90,87%	0,830
4	Naïve Bayes + PSO	88,34%	0,944

Dari hasil pengujian pada penelitian ini, dapat dilihat bahwa performansi akurasi Decision Tree C4.5 sangat meningkat ketika dilakukannya fitur seleksi oleh Algoritma Naïve Bayes menggunakan *Particle Swarm Optimization* (PSO) pada *dataset* gagal jantung dari nilai nilai *accuracy* 95,18 %, dan nilai AUC 0, 998. Kinerja algoritma pada penelitian ini menunjukkan performa yang sangat baik sekali dan tingkat akurasi yang sangat tinggi.

5. KESIMPULAN

Dari pengukuran kinerja dengan melakukan komparasi dua algoritma yang telah dilakukan berdasarkan jumlah data maka dapat disimpulkan bahwa algoritma *Decision Tree* memiliki kemampuan dalam pengambilan keputusan untuk menentukan penderita gagal Jantung. Adapun hasil penelitian dari percobaan yang telah dilakukan dapat disimpulkan bahwa:

1. Algoritma Decision Tree memiliki tingkat akurasi tertinggi yaitu 95,16% sedangkan Naïve Bayes sebesar 90,87%.
2. Model algoritma *Decision Tree* memiliki AUC sebesar 0,773 dan *Naïve Bayes* 0,830.
3. Algoritma Decision Tree C4.5 Akurasi meningkat ketika dilakukannya seleksi fitur PSO pada dataset prediksi penyakit gagal jantung dari nilai akurasi 0,95,16 % menjadi 95,18%.
4. AUC Algoritma Decision Tree C4.5 Nilai AUC juga mengalami peningkatan sebelum 0.773 menjadi AUC: 0.988.

DAFTAR PUSTAKA

- Cahya Putri Buani, Duwi. 2021. "Penerapan Algoritma Naïve Bayes Dengan Seleksi Fitur Algoritma Genetika Untuk Prediksi Gagal Jantung." *EVOLUSI : Jurnal Sains dan Manajemen* 9(2): 43–48.
- Derisma, D. 2020. "Perbandingan Kinerja Algoritma Untuk Prediksi Penyakit Jantung Dengan Teknik Data Mining." *Journal of Applied Informatics and Computing* 4(1): 84–88.
- Destiani, Dinda Karlia et al. 2018. "Klasifikasi Sinyal Ecgagal Jantung Menggunakan Wavelet Dan Jst Propagasi Balik Dengan Modifikasi Gradien Konjugat Polak-Ribiere Heart Failure Ecg Signal Classification Using Wavelet and Ann Backpropagation With Polak-Ribiere Conjugate Gradient." 5(1): 1811–18.
- Dharmawan, Weiskhy Steven. 2021. "I N F O R M a T I K a Dalam Prediksi Penyakit Jantung." *Jurnal Informatika, Manajemen dan Komputer* 13(2): 31–41.
- Effendi, Ferry, and Arief Muhamad. 2021. "Perbandingan Efektivitas Kombinasi Carvedilol – Ramipril Dan Bisoprolol – Candesartan Pada Pasien Gagal Jantung Kongestif Di Rsud Ciawi." *Jurnal Farmamedika (Pharmamedica Journal)* 6(1): 24–28.
- Hasanah, Quswatun, Hardian Oktavianto, and Yeni Dwi Rahayu. 2022. "Analisis Algoritma Gaussian Naive Bayes

- Terhadap Klasifikasi Data Pasien Penderita Gagal Jantung.” *Jurnal Smart Teknologi* 3(4): 382–89.
- hidayah nur umi, and dkk. 2021. “Hidayah Nur Umi Dkk.” *Nalisis Metode K Nearest Neighbor Terhadap Klasifikasi Data Pasien Penderita Gagal Jantung* 28: 145–58. <http://www.riss.kr/link?id=A99932365>.
- Informatika, Jurnal et al. 2022. “Penerapan Algoritma Naïve Bayes Untuk Mengetahui Pasien Penyakit Gagal Jantung Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM).” 2(September).
- Munandar, Tb Ai, and Agus Qomaruddin Munir. 2022. “Implementasi K-Nearest Neighbor Untuk Prototype Sistem Pakar Identifikasi Dini Penyakit Jantung.” *Respati* 17(2): 44.
- Nawawi, Hendri Mahmud, Jajang Jaya Purnama, and Agung Baitul Hikmah. 2019. “Komparasi Algoritma Neural Network Dan Naïve Bayes Untuk Memprediksi Penyakit Jantung.” *Jurnal Pilar Nusa Mandiri* 15(2): 189–94.
- Nurlia, Elin, and Ultach Enri. 2021. “Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5.” *Jurnal Teknik Informatika Musirawas)* Elin Nurlia 6(1): 42.
- Putra, Pandito Dewa, and Dian Palupi Rini. 2019. “Prediksi Penyakit Jantung Dengan Algoritma Klasifikasi.” *Prosiding Annual Research Seminar 2019* 5(1): 978–79.
- Putri, Trisca Anggraini, and Nurul Huda. 2020. “Analisis Dan Prediksi Penyakit Jantung Menggunakan Algoritma K-Means Clustering Pada Rumah Sakit Umum.” *Bina Darma Conference on Computer Science* 4(41): 197–206.
- Siahaan, Rahma Aulia, Marnis Nasution, and Mila Nirmala Sari Hasibuan. 2021. “Model Data Mining Untuk Perancangan Aplikasi Diagnostik Inflammatory Liver Disease.” *Jurnal Teknik Informatika UNIKA Santo Thomas* 06: 145–53.
- Subarkah, Pungkas, Wenti Risma, and Reza Aditya. 2022. “Comparison of Correlated Algorithm Accuracy Naive Bayes Classifier and Naive Bayes Classifier for Heart Failure Classification.” 14(2): 120–25.