

# PERBANDINGAN PREDIKSI PENYAKIT SERANGAN JANTUNG MENGGUNAKAN MODEL MACHINE LEARNING

Rivaldo Pranandito\*<sup>1)</sup>, Hendry<sup>2)</sup>

1. Universitas Kristen Satya Wacana, Salatiga, Indonesia
2. Universitas Kristen Satya Wacana, Salatiga, Indonesia

## Article Info

**Kata Kunci:** *Causes; Treatments; Heart Attack*

**Keywords:** *Causes; Treatments; Heart Attack*

## Article history:

Received 13 May 2023

Revised 27 May 2023

Accepted 10 June 2023

Available online 1 December 2023

## DOI :

<https://doi.org/10.29100/jipi.v8i4.4165>

\* Corresponding author.

Rivaldo Pranandito

E-mail address:

[672019320@student.uksw.edu](mailto:672019320@student.uksw.edu)

## ABSTRAK

Serangan jantung dapat disebabkan oleh tekanan darah, stres, terlalu banyak bekerja dan gula darah. Diantara organ tubuh manusia, jantung merupakan organ yang harus bekerja lebih keras memompa darah ke seluruh tubuh untuk menghasilkan oksigen dan nutrisi. Ketika jantung tidak berfungsi dengan baik, fungsi organ lain di dalam tubuh akan terganggu secara serius dan bahkan dapat menyebabkan gagal jantung. Naive Bayes merupakan probabilistik klasifikasi, yang berdasarkan teorema Bayes. Atribut fitur dapat diasumsikan dengan kontribusi yang setara ke dalam target. Dapat dikatakan bahwa regresi logistik adalah suatu teknik analisis data statistik yang bertujuan untuk mencari variabel yang saling berhubungan yang jawabannya bersifat kategorikal. Bagi sebagian orang, menentukan atribut target merupakan hal yang paling penting dalam klasifikasi data mining. Dalam pembuatan data langsung, semua variabel dianggap sebagai atribut independen. Hal ini, widget Selected Columns yang disertakan, harus memilih atribut target yang akan diperiksa. Dalam penelitian tidak mengatur jumlah metode yang digunakan. Oleh karena itu prediksi dan analisis dapat menyebabkan hasil yang kurang optimal. Oleh karena itu, metode dan perhitungan tambahan dapat dilakukan untuk penelitian lebih lanjut untuk mencapai hasil yang optimal dan serbaguna. Studi tambahan dan metode lain untuk akurasi yang lebih baik juga dapat dilakukan dengan informasi baru. Hasil penelitian dengan metode naïve bayes, logistic regression, dan neural network dihasilkan pencapaian terbaik untuk metode naïve bayes mendapatkan hasil akurasi 84,3%.

## ABSTRACT

Heart attacks can be caused by blood pressure, stress, overwork and blood sugar. Among the organs of the human body, the heart is an organ that has to work harder to pump blood throughout the body to produce oxygen and nutrients. When the heart is not functioning properly, the function of other organs in the body will be seriously disrupted and can even lead to heart failure. Naive Bayes is a probabilistic classification, which is based on Bayes' theorem. Feature attributes can be assumed to be equal contributions to the target. It can be said that logistic regression is a statistical data analysis technique that aims to find interrelated variables whose answers are categorical. For some people, determining the target attribute is the most important thing in data mining classification. In direct data generation, all variables are considered as independent attributes. This, the included Selected Columns widget, should select the target attribute to inspect. The research does not regulate the number of methods used. Therefore prediction and analysis can lead to sub-optimal results. Therefore, additional methods and calculations can be carried out for further research to achieve optimal and versatile results. Additional studies and other methods for better accuracy can also be carried out with the new information. The results of the research using the naïve Bayes, logistic regression, and neural network methods produced the best results for the logistic regression method to get an accuracy of 84.3%.

## I. PENDAHULUAN

**P**ENYAKIT jantung menyebabkan kematian karena tekanan darah, stres, terlalu banyak bekerja, gula darah dan banyak alasan lainnya. Jantung merupakan organ terpenting yang harus berfungsi dengan baik, karena tugasnya memompa darah ke seluruh tubuh agar oksigen dan nutrisi dapat disalurkan. Ketika jantung tidak berfungsi dengan baik, hal itu sangat mempengaruhi fungsi organ lain di dalam tubuh dan bahkan membuat Anda berisiko mengalami gagal jantung [1].

Dengan berkembangnya teknologi, banyak hal yang dapat dilakukan untuk meningkatkan kemudahan masyarakat, termasuk pengembangan bidang kecerdasan buatan. Kecerdasan buatan bisa digunakan di segala bidang, seperti computer vision yang bisa mendeteksi penyakit, atau sistem autopilot di lalu lintas. Maka, penelitian ini melihat ke bidang kecerdasan buatan yaitu machine learning yang dapat menggunakan informasi yang relevan untuk memprediksi kemungkinan penyakit jantung. Pada penelitian *Logistic Regression* hasil akurasi sebesar 81,7 %, 84,3 % untuk Naïve Bayes dan 78,2 % untuk Neural Network [2]. Tujuan dari penelitian ini untuk perbandingan prediksi bagaimana gejala awal untuk orang yang memiliki penyakit serangan jantung yang biasa disebut infark miokard akut (heart attack) merupakan penyakit kardiovaskular yang mematikan. Banyaknya kasus serta tingginya angka mortalitas akibat penyakit ini menyebabkan dibutuhkan suatu kajian yang menyeluruh dimana nantinya dapat dikembangkan lebih jauh lagi sehingga kasus serta tingkat mortalitasnya dapat menurun [3]. Maka dari itu pada penelitian tugas akhir ini penulis ingin membuat sebuah perbandingan yang berbeda dengan hasil penelitian sebelumnya menggunakan klasifikasi bagi penderita seseorang yang mengalami penyakit serangan jantung sederhana menggunakan model machine learning dalam penggunaan metode *Naïve Bayes*, *Logistic Regression* dan *Neural Network*. Penyakit kardiovaskular, atau yang biasa disebut penyakit jantung, umumnya mengacu pada kondisi di mana pembuluh darah menyempit atau tersumbat, yang dapat menyebabkan serangan jantung, angina, atau stroke. The American Heart Association menganggap bahwa kondisi jantung lain yang memengaruhi otot jantung, katup jantung, atau irama jantung sebagai bentuk penyakit jantung. Dalam istilah umum untuk gagal jantung dan tidak termasuk penyakit pembuluh darah yang menyebabkan penyakit jantung. Banyak orang menyebut penyakit jantung ini CVD. Menurut WHO (World Health Organization), penyakit jantung merupakan penyebab utama kematian di beberapa negara seperti Indonesia, Inggris Raya, Australia, Kanada, Amerika Serikat dan beberapa negara lainnya [4]. Banyak sekali jenis penyakit jantung, dan untuk mengetahui penyakit jantung tersebut, tidak ada salahnya mengikuti pembahasan singkat tentang penyakit jantung. Cardiac care in Cultural merupakan kegiatan dengan cara memberikan pelatihan manajemen serangan jantung seperti gejala serangan dan manajemen tindakannya dengan mempertahankan budaya tersebut selama tidak menyimpang dari Kesehatan [5]. Jantung merupakan Anggota tubuh empat bagian yang memompa darah kembali ke jantung dan ke seluruh tubuh. Akibatnya, hati memegang peranan penting sebagai penopang kehidupan manusia. Jika jantung bermasalah dengan pernafasan, juga dapat membahayakan kematian [3]. Penyakit jantung memang sangat mematikan, tetapi kewaspadaan dan juga pengetahuan mengenai penyakit jantung harus tetap dimiliki mulai umur sedini mungkin. Karena penyakit serangan jantung berawal dari kelalaian hidup ketika masih muda [6]. Diharapkan pasien serangan jantung dapat mempertahankan dan menstabilkan fungsi tubuh secara optimal, mencapai kesehatan yang optimal, mencegah kekambuhan dan meningkatkan kualitas hidup. Kepatuhan minum obat (pemeriksaan rutin), penggunaan obat resep, kepatuhan diet yang diresepkan dan olahraga yang cukup. Berdasarkan penelitian yang telah dilakukan oleh di RSUP Dr Kariadi Semarang dengan total sampel 128 jiwa didapatkan hasil 80,50% dengan jumlah pasien 103 berpotensi mengidap penyakit jantung [7]. Penelitian oleh Nisa Trianifa menunjukkan bahwa metode SVM dengan ekstraksi sinyal EKG menggunakan algoritma pembacaan waktu dan tekstur sinyal memiliki kinerja yang baik dalam mengklasifikasikan penyakit jantung. Hasilnya mencapai akurasi 95%, sensitivitas 100%, spesifitas 91,67%, dan presisi 100%. Metode ini memiliki potensi sebagai alat bantu penting dalam diagnosis penyakit jantung, tetapi perlu dilakukan validasi lebih lanjut sebelum diterapkan secara luas dalam praktik klinis. Berdasarkan penelitian terkait, peneliti akan melakukan penelitian untuk melakukan prediksi penyakit jantung yang akan dibandingkan dengan 3 metode yaitu logistic regression, naïve bayes, dan neural network [8].

## II. METODE PENELITIAN

Penelitian ini menggunakan metode perbandingan prediksi Naive Bayes, Logistic Regression, Neural Network untuk menganalisa prediksi akurasi terbaik, berdasarkan data serangan jantung Kaggle, yang terkait dengan Kesehatan jantung seseorang. Tujuan utama dari penelitian untuk mengetahui akurasi terbaik dalam mengetahui prediksi yang muncul dalam dataset yang digunakan. Dalam penelitian ini, peneliti menggunakan fakta serangan jantung yang menjelaskan bagaimana masalah yang diteliti dapat dipecahkan. Untuk pengumpulan data dari Kaggle diolah menggunakan data sampler aplikasi orange yang telah dipersiapkan sebelumnya [9].

### A. Machine Learning

Machine learning atau kecerdasan buatan merupakan rancangan sistem yang dapat membuat prediksi berdasarkan pengalaman. Metode ini menggunakan data training untuk membuat model yang memberikan informasi baru sebagai masukan, dapat digunakan untuk memprediksi penyakit jantung. Dengan model ini, dapat menemukan pola data dalam input dataset untuk membangun model dan membuat prediksi yang akurat dari data baru. Sebelumnya, kumpulan data dibersihkan dan nilai yang hilang ditangani. Tes akurasi kemudian dilakukan untuk memprediksi penyakit jantung menggunakan data input baru. Secara definisi, machine learning dipakai dengan penggunaan AI yang dibuktikan dengan algoritma yang berjalan di komputer. Martin Molina (2020) mendefinisikan Intelligent System sebagai berikut [10]: “Sebuah sistem yang:

- beroperasi pada suatu lingkungan bersama agen,
- memiliki kemampuan kognitif seperti mengerti, mengontrol tindakan, bernalar, maupun berbahasa,
- bertindak secara rasional, serta
- memiliki kemampuan untuk beradaptasi dengan belajar.”

### B. Naive Bayes

Naive Bayes merupakan metode klasifikasi probabilistik yang berdasarkan teorema Bayes, mengasumsikan bahwa dalam setiap fitur secara independen mempengaruhi beberapa kelas target. Pengklasifikasian Naive Bayes mudah dalam implemetasi, cepat dalam komputasi, dan bekerja dengan baik untuk kumpulan big data dan berdimensi tinggi. Metode ini cocok dalam aplikasi real-time, tidak terlalu sensitif terhadap noise. Pengklasifikasi metode Naive Bayes memproses kumpulan data training untuk menghitung perkiraan probabilitas kelas dan probabilitas bersyarat yang menentukan frekuensi setiap nilai. Dalam fitur untuk nilai kelas tertentu dibagi dengan frekuensi nilai kelas tersebut. Dalam Persamaan (1) dan Persamaan (2), teorema Bayes menyediakan cara untuk menghitung probabilitas posterior  $P(c|x)$  menggunakan  $P(c)$ ,  $P(x)$ , dan  $P(x|c)$  untuk menghitung. Pengklasifikasian metode Naive Bayes dapat diasumsikan bahwa efek dari nilai prediktor ( $x$ ) pada kelas tertentu ( $c$ ) tidak bergantung pada nilai prediktor lainnya. Asumsi ini disebut independensi bersyarat kelas.  $P(c | x) = P(x|c)P(c) / P(x)$  (1)  $P(c | X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$  (2) •  $P(c|x)$  adalah probabilitas kelas posterior (target) yang diberikan prediksi (atribut). •  $P(c)$  adalah probabilitas dari kelas sebelumnya. •  $P(x|c)$  adalah probabilitas, jadi probabilitas kelas yang diberikan oleh prediktor. •  $P(x)$  adalah probabilitas awal dari prediktor. Berdasarkan penjelasan diatas bahwa dari rumus tersebut dapat menyebabkan semakin banyak faktor syarat yang mempengaruhi nilai probabilitas yang mengakibatkan untuk perhitungan, prediksi dan analisis [11].

### C. Logistic Regression

Regresi logistik adalah analisis data dalam statistik yang bertujuan untuk menentukan hubungan antara beberapa variabel ketika variabel responsnya kategoris baik dalam nilai nominal maupun ordinal dan variabel penjelasnya adalah metode kategoris atau kontinu. Metode dan prosedur regresi logistik hampir identik dengan metode regresi linier. Model yang digunakan untuk menentukan persamaan logistik adalah:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}}$$

Jadi persamaan logistiknya adalah:

$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \ln \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}}$  Logistic Regression digunakan untuk mengevaluasi faktor-faktor yang terkait pengetahuan dan pemahaman yang buruk tentang penyakit kardiovaskular (stroke dan jantung gejala serangan) [12].

### D. Neural Network

Neural network (NN) merupakan metode yang terinspirasi oleh jaringan saraf, di mana sel saraf (neuron) saling terhubung. Metode yang sering digunakan dalam NN adalah backpropagation (BP), yang merupakan metode

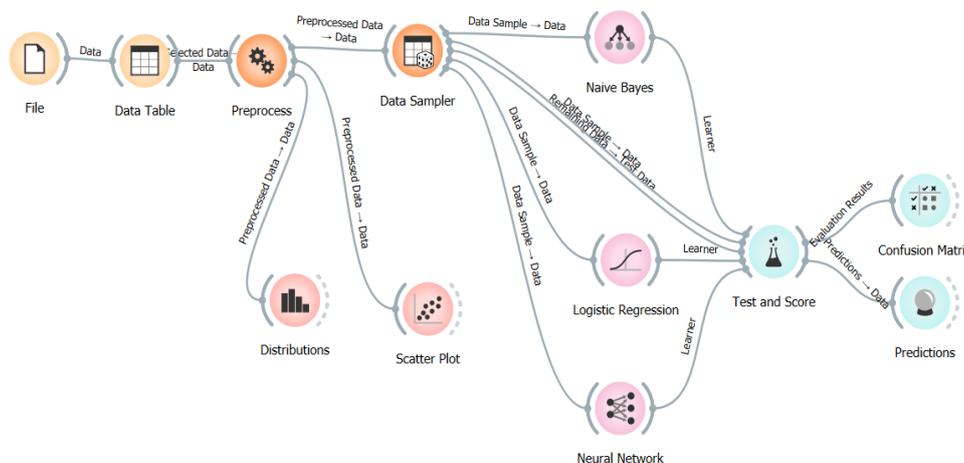
sistematis. Neural network memiliki dasar matematika yang bersifat objektif, dan algoritmanya mempertahankan bentuk persamaan dan koefisien dalam rumus dengan mengurangi jumlah kesalahan kuadrat. Tujuan dari metode ini adalah untuk mengidentifikasi pola pada data historis dan kemudian menggeneralisasi pola tersebut, sehingga dapat mengikuti pola data atau kurva. Algoritma dan parameter yang tepat sangat penting untuk menentukan akurasi, termasuk dalam menentukan jumlah data yang diperlukan [13]. Untuk perancangan system prediksi penyakit serangan jantung, dalam penelitian ini telah digunakan algoritma CNN (Convolutional Neural Network) [14].

### III. HASIL DAN PEMBAHASAN

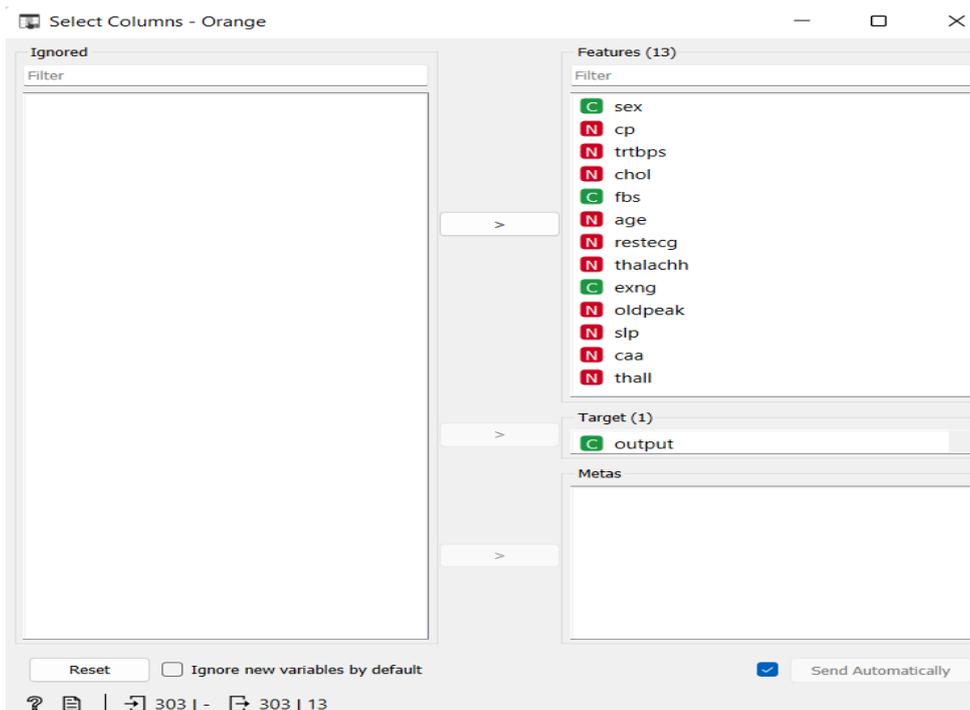
Pada penelitian serangan jantung ini terdapat berbagai simpulan dari model “*Naïve Bayes, Logistic Regression, dan Neural Network*”, yang terwujud. Hasil dan pembahasan dijelaskan sebagai berikut.

#### A. Pemetaan Atribut

Terlihat ada tiga bagian yaitu kuning, pink dan hijau. Dalam bagian kuning adalah tahap dalam persiapan data, yaitu. impor file data, pilih kolom untuk menentukan atribut gratis dan atribut target, lalu gunakan tabel data untuk membuat tampilan tabel. Bagian merah menunjukkan metode pembelajaran terbimbing yang diklasifikasikan. Penelitian ini menggunakan tiga metode untuk membandingkan data penilaian. Bagian hijau merupakan hasil evaluasi. Dalam menentukan atribut target adalah hal terpenting dalam klasifikasi di data mining. Saat menghasilkan data secara langsung, semua variabel dianggap sebagai atribut independen. Dalam hal ini, widget Selected Columns harus disertakan untuk pemilihan atribut target yang akan diperiksa. Sebelum melakukan analisis teks, langkah awal yang dilakukan adalah *preprocess text*. Pada tahap ini, teks dipisahkan menjadi unit-unit yang lebih kecil yang disebut token, kemudian dilakukan *Filtering* sebanyak 80% dari data asli. Urutan langkah-langkah ini dapat diaktifkan atau dinonaktifkan menggunakan *widget preprocess text* di Orange Data Mining [15]. Dalam data ini menggunakan atribut : sex, cp, trtbps, chol, fbs, age, restecg, thalachh, exng, oldpeak, slp, caa, thall, output(target).



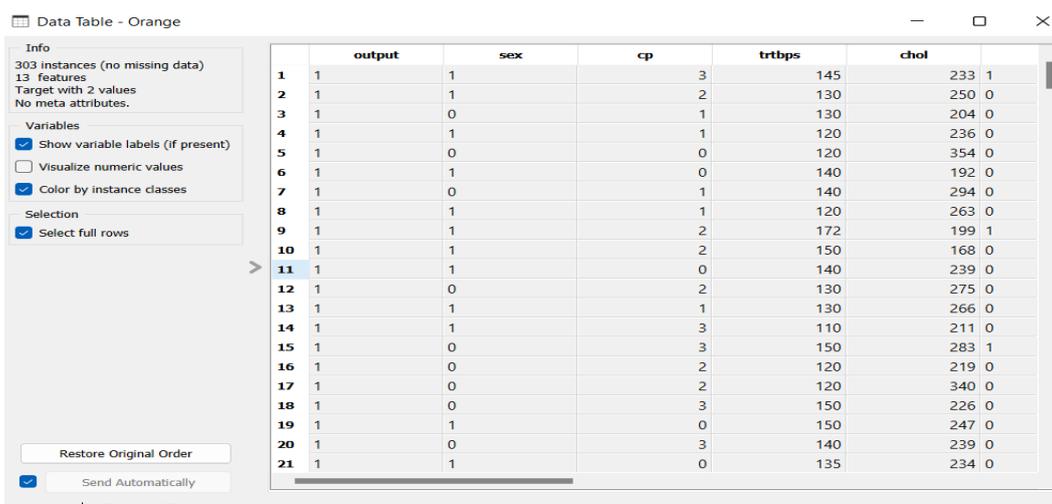
Gambar. 1. Workflow penelitian dengan software Orange



Gambar 2. Pemetaan atribut bebas dan target

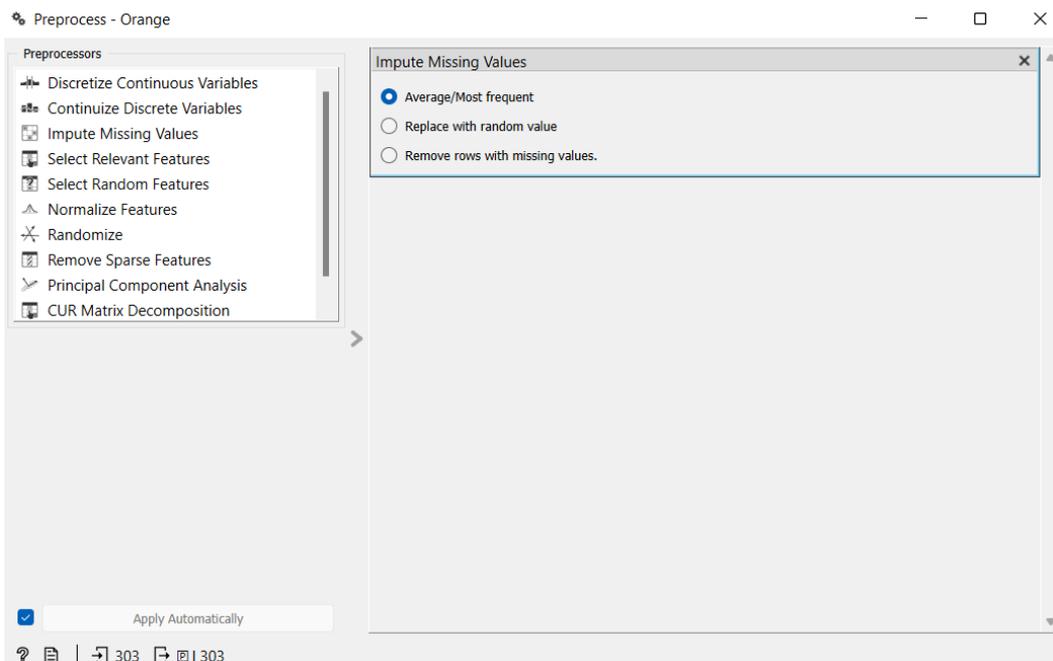
### B. Pembagian Data

Penelitian ini menggunakan widget pengujian dan evaluasi untuk mendistribusikan data pelatihan dan pengujian. Menurut tabel data, filter diberi perlakuan awal dengan 80% data sampel. Hasil dari confusion matrix, skor rating, tes, dan widget rating dalam confusion matrix, untuk kedua metode yang digunakan dalam penelitian ini. Nilai prediktif yang benar adalah kondisi dimana terdapat hasil positif yang benar (di mana hasil 1 memprediksi serangan jantung) dan hasil negatif yang sebenarnya (dimana hasil 0 tidak memprediksi serangan jantung). Jumlah total hasil positif sejati dan negatif sejati untuk metode regresi logistik, naif Bayes, dan jaringan saraf masing-masing adalah 197, 205, dan 190. Hasil klasifikasi yang benar dalam metode Naive Bayes adalah yang terbaik, diikuti oleh Logistic Regression. regresi. Berdasarkan confusion matrix, akurasi dan tingkat akurasi masing-masing metode dapat dihitung seperti yang ditunjukkan pada gambar. Hasil yang akan diprediksi dalam penelitian merupakan prediksi akurasi terbaik dalam penentuan confusion matrix, dan evaluasi modelling yang sudah ditentukan.

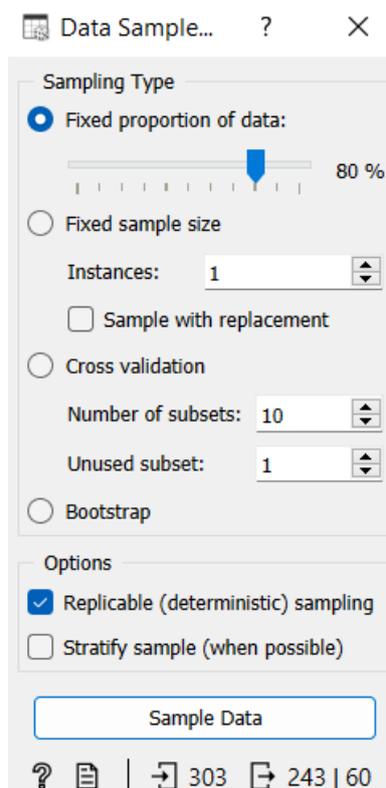


	output	sex	cp	trtbps	chol	
1	1	1	3	145	233	1
2	1	1	2	130	250	0
3	1	0	1	130	204	0
4	1	1	1	120	236	0
5	1	0	0	120	354	0
6	1	1	0	140	192	0
7	1	0	1	140	294	0
8	1	1	1	120	263	0
9	1	1	2	172	199	1
10	1	1	2	150	168	0
11	1	1	0	140	239	0
12	1	0	2	130	275	0
13	1	1	1	130	266	0
14	1	1	3	110	211	0
15	1	0	3	150	283	1
16	1	0	2	120	219	0
17	1	0	2	120	340	0
18	1	0	3	150	226	0
19	1	1	0	150	247	0
20	1	0	3	140	239	0
21	1	1	0	135	234	0

Gambar 3 Data Table Spreadsheet



Gambar. 4. *Preprocess Average*



Gambar. 5. *Pemilihan proportion dari data sampler*

### C. Confusion Matrix

Setelah melakukan pengujian data maka selanjutnya melakukan pengujian evaluasi performa menggunakan confusion matrix untuk mengukur performa berupa akurasi, presisi, dan recall. Hasil klasifikasi menggunakan metode *Naïve Bayes*, *Logistic Regression*, dan *Neural Network* sehingga menghasilkan perhitungan yang telah diklasifikasi seperti ditunjukkan pada table dibawah.

		Predicted		$\Sigma$
		0	1	
Logistic Regression	0	78	31	109
	1	15	119	134
	$\Sigma$	93	150	243

		Predicted		$\Sigma$
		0	1	
Naive Bayes	0	90	19	109
	1	19	115	134
	$\Sigma$	109	134	243

		Predicted		$\Sigma$
		0	1	
Neural Network	0	79	30	109
	1	23	111	134
	$\Sigma$	102	141	243

Gambar 6 Confusion Matrix

Setelah mendapatkan hasil dari Confusion Matrix maka dapat di simpulkan bahwa:

- Logistic Regression menghasilkan True Postive (TP) : 78, False Positive (FP) : 31, False Negative (FN) : 15, dan True Negative (TN) : 119. Dengan total keseluruhan data (n) : 243.
- Naïve Bayes menghasilkan True Postive (TP) : 90, False Positive (FP) : 19, False Negative (FN) : 19, dan True Negative (TN) : 115. Dengan total keseluruhan data (n) : 243.
- Neural Network menghasilkan True Postive (TP) : 79, False Positive (FP) : 30, False Negative (FN) : 23, dan True Negative (TN) : 111. Dengan total keseluruhan data (n) : 243.

#### D. Evaluasi Model

Dapat dilihat bahwa model yang paling akurat adalah model Naive Bayes, diikuti oleh regresi logistik dan jaringan saraf tingkat terendah. Berbeda dalam pengukuran tingkat akurasi sesuai dengan kemampuan model untuk membuat prediksi yang benar, tingkat akurasi disebut nilai prediksi positif, yaitu rasio hasil tes positif. Dapat dilihat bahwa Naive Bayes adalah model level paling akurat, diikuti oleh regresi logistik dan jaringan saraf level terendah. Gambar 8 menunjukkan evaluasi kedua model ini berdasarkan pelatihan dan pengujian.

TABEL I  
EVALUASI MODEL

Model	AUC	CA	F1	PRECISION	RECAL	TRAIN TEST [S]	TEST TIME [S]
Neural Network	0.868	0.782	0.781	0.782	0.782	1.107	0.025
Naïve Bayes	0.897	0.844	0.844	0.844	0.844	0.041	0.009
Logistic Regression	0.871	0.811	0.809	0.814	0.811	0.518	0.015

#### E. Accuracy

Jadi pada pengujian akurasi yang diperoleh dengan menggunakan 80% data sampler menghasilkan akurasi Logistic Regression 81,7%, Naïve Bayes 84,3%, dan recall 50%, yang didapatkan dari table confusion matrix.

- Logistic Regression :
  - True Postive (TP) : 78

- False Positive (FP) : 31
- False Negative (FN) : 15
- True Negative (TN) : 119

$$\begin{aligned} \text{Accuracy} &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \\ &= \frac{(78 + 119)}{(78 + 31 + 15 + 119)} \\ &= 0.817 \end{aligned}$$

$$\begin{aligned} \text{Accuracy Logistic Regression} &= 0.817 \times 100\% \\ &= 81.7\% \end{aligned}$$

- Naïve Bayes :
  - True Postive (TP) : 90
  - False Positive (FP) : 19
  - False Negative (FN) : 19
  - True Negative (TN) : 115

$$\begin{aligned} \text{Accuracy} &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \\ &= \frac{(90 + 115)}{(90 + 19 + 19 + 115)} \\ &= 0.843 \end{aligned}$$

$$\begin{aligned} \text{Accuracy Naive Bayes} &= 0.843 \times 100\% \\ &= 84.3\% \end{aligned}$$

- Neural Network
  - True Postive (TP) : 79
  - False Positive (FP) : 30
  - False Negative (FN) : 23
  - True Negative (TN) : 111

$$\begin{aligned} \text{Accuracy} &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \\ &= \frac{(79 + 111)}{(79 + 30 + 23 + 111)} \\ &= 0.782 \end{aligned}$$

$$\begin{aligned} \text{Accuracy Neural Network} &= 0.782 \times 100\% \\ &= 78.2\% \end{aligned}$$

TABEL II  
PERBANDINGAN AKURASI

Model	Akurasi
<i>Logistic Regression</i>	81.7%
<i>Naïve Bayes</i>	84.3%
<i>Neural Network</i>	78.2%

Pada penelitian ini dapat digunakan di semua area, namun algoritma yang ada dapat menghasilkan informasi baru. Algoritma untuk pembelajaran mesin terus dikembangkan dalam beberapa penelitian. Tren terkini menunjukkan bahwa algoritma pembelajaran mesin telah dikembangkan di bidang medis. Hasil dari penelitian ini menunjukkan bahwa penggunaan machine learning dengan tools orange dalam pengembangan model, dapat meningkatkan kemudahan *user* dalam mereka melakukan penelitian kuantitatif. Data dari metode naïve bayes diprediksi tepat waktu dengan Train Test 0.041 detik, untuk Test Time dengan catatan waktu 0.009 detik, logistic regression dengan Train Test 0.518, untuk Test Time dengan catatan waktu 0.015, dan neural network mendapatkan catatan waktu Train Test 1.107, kemudian Test Time 0.025. Nilai precision, recall, dan f1-score diperoleh dari tabel confusion matrix yang terbentuk pada masing- masing fold [16]. Melalui penelitian sebelumnya ditemukan bahwa metode yang digunakan hanya menggunakan naïve bayes. Sedangkan untuk penelitian sekarang dilakukannya uji coba data dengan metode logistic regression, naïve bayes, dan neural network.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian mendapatkan perbandingan antara metode yang paling baik untuk digunakan dalam hal prediksi bagi penyakit serangan jantung merupakan metode Naive Bayes. Dan hasil dari sebuah Penelitian menunjukkan bahwa menggunakan aplikasi Orange untuk pengembangan data dapat menghemat waktu, karena pengguna hanya perlu membuat alur kerja dan mengonfigurasinya dengan satu klik, tanpa pemrograman Python. Pengguna non-IT, terutama dalam penelitian kesehatan, pertanian, dan bidang lainnya, dapat menggunakan perangkat lunak ini untuk mengklasifikasikan data mining. Hasil menunjukkan bahwa pelatihan dan pengujian model membutuhkan waktu yang relatif singkat. Hasil dari penelitian serangan jantung menggunakan data hearth attack yang bertujuan untuk mendapatkan perbandingan antara prediksi dan analisis seseorang terkena serangan jantung atau tidak. Dapat disimpulkan bahwa Naive Bayes merupakan metode terbaik untuk memprediksi dan menganalisa seseorang terkena serangan jantung dengan score akurasi yang tinggi 84,3 % , Logistic Regression mendapatkan score 81,7 % sedangkan Neural Network mendapatkan score 78,2 % . Metode Naive Bayes membutuhkan waktu training, testing lebih cepat dari pada Logistic Regression dan Neural Network. Penelitian ini tidak mengatur banyak metode yang digunakan. Sehingga saat melakukan prediksi dan analisis mungkin mendapatkan hasil yang kurang maksimal. Oleh karena itu, untuk penelitian selanjutnya dapat dilakukan penambahan metode dan perhitungan sehingga menghasilkan hasil yang optimal dan beragam. Selain itu, dengan informasi baru, penyelidikan lebih lanjut dapat dilakukan dan mungkin ada metode lain untuk mencapai akurasi yang lebih baik.

#### V. DAFTAR PUSTAKA

- [1] S. Dwi Andini, "Review Analisis Hubungan Penyakit Jantung Koroner Terhadap Risiko Stres," *Cerdika J. Ilm. Indones.*, vol. 2, no. 11, pp. 933–937, 2022, doi: 10.36418/cerdika.v2i11.471.
- [2] S. A. T. Al Azhima, D. Darmawan, N. F. A. Hakim, I. Kustiawan, M. Al Qibtiya, and N. S. Syafei, "Hybrid Machine Learning Model untuk Memprediksi Penyakit Jantung dengan Metode Logistic Regression dan Random Forest," *J. Teknol. Terpadu*, vol. 8, no. 1, pp. 40–46, 2022.
- [3] T. Aniamarta, A. Salsabilla Huda, and F. Lizariani Aqsha, "Review Artikel: Penyebab dan Pengobatan Serangan Jantung Review Article: Causes and Treatments of Heart Attack," *Biol. Samudra*, vol. 4, no. 1, pp. 22–31, 2022.
- [4] American Journal of Sociology, "Penyakit Jantung," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [5] M. S. Mohtar, M. Riduansyah, S. Suwardi, and A. Asmadiannor, "Cardiac Care in Cultural (Ccic): Pelatihan Manajemen Serangan Jantung Dengan Budaya Banjar 'Menyamak' Bagi Ahli Cabut Angin Di Desa Pemurus Baru Banjarmasin," *J. Suaka Insa. Mengabdi*, vol. 3, no. 2, pp. 123–132, 2022, doi: 10.51143/jsim.v3i2.312.
- [6] W. N. Santosa and B. Baharuddin, "Penyakit Jantung Koroner dan Antioksidan," *KELUWIH J. Kesehat. dan Kedokt.*, vol. 1, no. 2, pp. 98–103, 2020, doi: 10.24123/kesdok.v1i2.2566.
- [7] T. Tajudin, I. D. W. Nugroho, and V. Faradiba, "ANALISIS KOMBINASI PENGGUNAAN OBAT PADA PASIEN JANTUNG KORONER (Coronary Heart Disease) DENGAN PENYAKIT PENYERTA DI RUMAH SAKIT X CILACAP TAHUN 2019," *Pharmaqueous J. Ilm. Kefarmasian*, vol. 1, no. 2, pp. 6–13, 2020, doi: 10.36760/jp.v1i2.111.
- [8] B. Penelitian *et al.*, "Klasifikasi Penyakit Jantung Menggunakan Metode Support Vector Machine Berdasarkan Perbandingan Algoritma Pembacaan Waktu Dengan Tekstur Sinyal Sebagai Metode Ekstraksi Sinyal Ekg," vol. 02, no. 01, pp. 7–11, 2020.
- [9] I. R. Hikmah and R. N. Yasa, "Perbandingan Hasil Prediksi Diagnosis pada Indian Liver Patient Dataset (ILPD) dengan Teknik Supervised Learning Menggunakan Software Orange," *J. Telemat.*, vol. 16, no. 2, pp. 69–76, 2021.
- [10] F. Chollet, "Machine learning 분야 소개 및 주요 방법론 학습 기본 machine learning 알고리즘에 대한 이해 및 응용 관련 최신 연구 동향 습득," *Mach. Learn.*, vol. 45, no. 13, pp. 40–48, 2017, [Online]. Available: <https://books.google.ca/books?id=EoYBngEACAAJ&dq=mitchell+machine+learning+1997&hl=en&sa=X&ved=0ahUKewiomdqfj8TkAhWGslkKHRCbAtoQ6AEIKjAA>.
- [11] W. E. Nugroho, A. Sofyan, and O. Somantri, "Metode Naive Bayes Dalam Menentukan Program Studi Bagi Calon Mahasiswa Baru," *Infotekmesin*, vol. 12, no. 1, pp. 59–64, 2021, doi: 10.35970/infotekmesin.v12i1.491.
- [12] S. Annas, A. Aswi, M. Abdy, and B. Poerwanto, "Binary Logistic Regression Model of Stroke Patients: A Case Study of Stroke Centre Hospital in Makassar," *Indones. J. Stat. Its Appl.*, vol. 6, no. 1, pp. 161–169, 2022, doi: 10.29244/ijsa.v6i1p161-169.

- [13] Y. Novi, “Penerapan Metode Neural Network dengan Struktur Backpropagation untuk Prediksi Stok Obat di Apotik,” *Semin. Nas. Apl. Teknol. Inf.*, vol. 2011, no. Snati, pp. 17–18, 2011.
- [14] Ajib Susanto, Yupie Kusumawati, Ericsson Dhimas Niagara, and Christy Atika Sari, “Convolutional Neural Network Dalam Sistem Deteksi Helm Pada Pengendara Motor,” *Semin. Nas. Teknol. dan Multidisiplin Ilmu*, vol. 2, no. 1, pp. 91–99, 2022, doi: 10.51903/semnastekmu.v2i1.158.
- [15] A. Patunru, A. Surianta, and A. Tenggara, “Menarik Penanaman Modal Asing ( PMA ),” no. 4, pp. 1–9, 2020.
- [16] I. W. Saputro and B. W. Sari, “Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa,” *Creat. Inf. Technol. J.*, vol. 6, no. 1, p. 1, 2020, doi: 10.24076/citec.2019v6i1.178.