

**PREDIKSI GAGAL JANTUNG DENGAN MENGGUNAKAN  
METODE C4.5 DAN NAÏVE BAYES**

**PROPOSAL TUGAS AKHIR**



Diajukan Oleh :

Julia Triani

8020190232

Untuk Persyaratan Penelitian Dan Penulisan Tugas Akhir

Sebagai Akhir Proses Studi Strata 1

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS DINAMIKA BANGSA**

**2022**

## **IDENTITAS PROPOSAL PENELITIAN**

Judul Proposal : PREDIKSI GAGAL JANTUNG DENGAN  
MENGUNAKAN METODE C4.5 DAN NAÏVE BAYES

Program Studi : Teknik Informatika

Jenjang Pendidikan : Strata 1 (S1)

Peneliti :

- a. Nama Lengkap : Julia Triani
- b. NIM : 8020190232
- c. Jenis Kelamin : Perempuan
- d. Tempat/Tgl. Lahir : Gardu Harapan / 27 Juli 2001
- e. Alamat : Air Hitam Rt 01, Desa Kebon IX  
Kab. Muaro Jambi – Sungai  
Gelam, Jambi 36374
- f. No. Telepon : 082175871772
- g. E-mail : juliatriani2707@gmail.com

## PERNYATAAN HASIL EVALUASI

Nama : Julia Triani  
NIM : 8020190232  
Prodi : Teknik Informatika  
Judul : PREDIKSI GAGAL JANTUNG DENGAN MENGGUNAKAN  
METODE C4.5 DAN NAÏVE BAYES

1. Hasil Evaluasi : Disetujui / Disetujui dengan perbaikan / Ditolak \*)

2. Alasan Penolakan Skripsi :

- Proyek Skripsi tidak relevan dengan Program Studi
- Pernah ada topik sejenis
- Metode utama telah banyak dipakai
- Metode yang dipakai tidak jelas
- Masalah terlalu sempit
- .....

3. Catatan Perbaikan :

.....  
.....  
.....  
.....

Mengetahui,  
Ketua Prodi Teknik Informatika

Beny S.Kom, M.Sc

## LEMBAR PERSETUJUAN PROPOSAL TUGAS AKHIR

Judul Proposal : PREDIKSI GAGAL JANTUNG DENGAN  
MENGUNAKAN METODE C4.5 DAN NAÏVE BAYES

Nama : Julia Triani

NIM : 8020190232

Program Studi : Teknik Informatika

Menyetujui,

Pembimbing I

Pembimbing II

(-----)

(-----)

Ketua Program Studi

(Beny S.Kom, M.Sc)

NIK : YDB.07.84.043

## 1. LATAR BELAKANG

Pada masa sekarang ini kebutuhan terhadap analisis data sangat penting dan dibutuhkan. Kemajuan perkembangan data yang semakin pesat mendorong untuk memanfaatkan data dalam pencarian informasi maupun pengetahuan. Salah satu cara pendekatan yang bisa dilakukan untuk menganalisis sekumpulan data yaitu dengan cara mengklasifikasikan data tersebut. Klasifikasi merupakan salah satu teknik data mining yang digunakan untuk membangun suatu model dari sampel data yang belum terklasifikasi untuk digunakan serta mengklasifikasi sampel data baru ke dalam kelas-kelas yang sejenis atau kelas-kelas yang sama.

Industri kesehatan memiliki data kesehatan yang cukup besar, hanya saja sebagian besar data tersebut tidak diolah untuk mengetahui informasi tersembunyi untuk dijadikan pengambilan keputusan yang efektif oleh para praktisi kesehatan. Pengambilan keputusan atas dasar data dan informasi yang akurat akan menghasilkan keputusan dan prediksi penyakit menjadi tepat sasaran. Penyakit jantung di Indonesia merupakan penyakit nomor satu yang mendorong angka kematian yang cukup tinggi. Oleh sebab itu penyakit jantung perlu diprediksi lebih lanjut dengan menggunakan klasifikasi data mining dengan tujuan agar praktisi kesehatan dalam pengambilan keputusan bisa lebih tepat dan akurat [1].

Pada tahun 2005 sedikitnya 17,5 juta atau setara dengan 30% kematian di seluruh dunia disebabkan oleh penyakit jantung. Faktor gejala yang terdiagnosa sebagai penyakit jantung antara lain adalah jenis sakit dada (*chest pain*), tekanan darah tinggi (*tresbps*), kolesterol (*chol*), nilai tes EKG (*resting electrodiagraphic "restacg"*), denyut jantung (*thalach*) dan kadar gula (*fasting blood sugar "FBS"*), dan beberapa faktor lainnya yang mengidentifikasi bahwa seseorang mempunyai penyakit jantung (Rifai, 2013). Menurut Han dan Kamber dalam (Arisandy, 2017), klasifikasi *Naïve Bayes* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam basis data dengan jumlah yang besar. Penelitian yang dilakukan oleh Heru Sulistiono (2015) berjudul "*Kajian Penerapan Algoritma C4.5, Neural Network Dan Naïve Bayes Untuk Klasifikasi Mahasiswa Yang Bermasalah Dalam Registrasi*". Algoritma *Naive*

*Bayes* memiliki tingkat akurasi yang paling tinggi dibandingkan dengan algoritma *Neural Network* dan *C4.5* dengan nilai persentase akurasi 93,58% [2].

Penelitian yang dilakukan oleh Anbarasi dkk (2010) dalam memprediksi kelangsungan hidup penyakit jantung dengan berdasarkan 909 kasus dan 6 Atribut dengan menggunakan metode *Naïve Bayes*, *Decision Tree* dan *Clasification Via Clustering*. Hasil penelitian tersebut metode *Decision Tree* menghasilkan nilai terbaik.

Berdasarkan penelitian diatas, maka penulis menguraikan penelitian dengan judul “ **PREDIKSI GAGAL JANTUNG DENGAN MENGGUNAKAN METODE C4.5 DAN NAÏVE BAYES**” peneliti akan memilih metode *C4.5* dan *Naïve Bayes* untuk memprediksi gagal jantung serta mengukur tingkat akurasi dalam penyakit jantung sehingga terbentuk modelnya, dengan mengoptimal atribut yang berasal dari dataset yang terpercaya untuk memprediksi penyakit jantung dengan tujuan agar akurasi menjadi meningkat.

## **2. RUMUSAN MASALAH**

Berdasarkan latar belakang yang telah di uraikan di atas, maka yang menjadi rumusan masalah dalam penelitian ini adalah bagaimana memprediksi serta mengukur tingkat akurasi penyakit gagal jantung menggunakan metode *C4.5* dan *Naïve Bayes*.

## **3. BATASAN MASALAH**

Agar tidak menyimpang dalam penulisan tugas akhir ini dari permasalahan dan lebih terarah, maka perlu adanya batasan masalah pada ruang lingkup bahasan antara lain sebagai berikut :

4. Menggunakan dan membahas algoritma klasifikasi *C4.5* dan *Naïve Bayes* dalam memprediksi penyakit gagal jantung.
5. Aplikasi yang akan digunakan untuk memproses data adalah *tools Rapid Miner*.
6. *Dataset* prediksi gagal jantung yang akan digunakan dalam penelitian ini diambil pada laman <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>

## **4. TUJUAN DAN MANFAAT PENELITIAN**

### **4.1 TUJUAN PENELITIAN**

Dari rumusan masalah yang terjadi maka penelitian ini bertujuan untuk:

1. Mengetahui bagaimana algoritma klasifikasi *C4.5* dan *Naïve Bayes* memprediksi penyakit gagal jantung.
2. Mengetahui perbandingan antara algoritma klasifikasi *C4.5* dan *Naïve Bayes* dalam mengukur tingkat akurasi penyakit gagal jantung.

### **4.2 MANFAAT PENELITIAN**

Adapun manfaat yang terdapat dalam penelitian ini adalah :

1. Penulis dapat menambah pengetahuan terkait bagaimana cara mencari informasi yang penting dalam suatu data dengan menggunakan metode Algoritma Klasifikasi *C4.5* dan *Naïve Bayes*
2. Pengembangan Data Mining dalam dunia kesehatan.
3. Menjadi bahan pertimbangan kesehatan dalam perawatan pasien gagal jantung.
4. Dapat digunakan sebagai acuan untuk melakukan penelitian berikutnya.

## **5. LANDASAN TEORI**

### **5.1 PENYAKIT GAGAL JANTUNG**

Gagal jantung adalah keadaan dimana jantung tidak mampu memompakan darah dengan jumlah yang sesuai untuk memenuhi kebutuhan metabolik tubuh, atau kemampuan tersebut yang hanya dapat dicapai dengan tekanan pengisian jantung yang tinggi, atau keduanya. Definisi lain ialah sindroma kompleks sebagai akibat dari kelainan jantung secara fungsional maupun struktural yang mengganggu kemampuan jantung sebagai pompa untuk mendukung sirkulasi fisiologis. Sindroma dari gagal jantung dicirikan oleh gejala-gejala seperti mudah lelah dan sesak nafas, dan tanda-tanda seperti retensi cairan.

Gagal jantung dapat disebabkan oleh banyak hal. Secara epidemiologi cukup penting untuk mengetahui penyebab dari gagal jantung, di negara berkembang penyakit

arteri koroner dan hipertensi merupakan penyebab terbanyak sedangkan di negara berkembang yang menjadi penyebab terbanyak adalah penyakit jantung katup dan penyakit jantung akibat malnutrisi [3].

## **5.2 DATA MINING**

Menurut Rusdianto dkk [4] “Data mining dapat diartikan sebagai menambang data atau upaya untuk menggali informasi yang berharga dan berguna pada *database* yang sangat besar”.

Menurut Yuli Mardi [5] “Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu”.

Pengertian data mining diatas dapat disimpulkan bahwa data mining merupakan proses mencari data atau informasi yang terpilih menggunakan metode-metode tertentu.

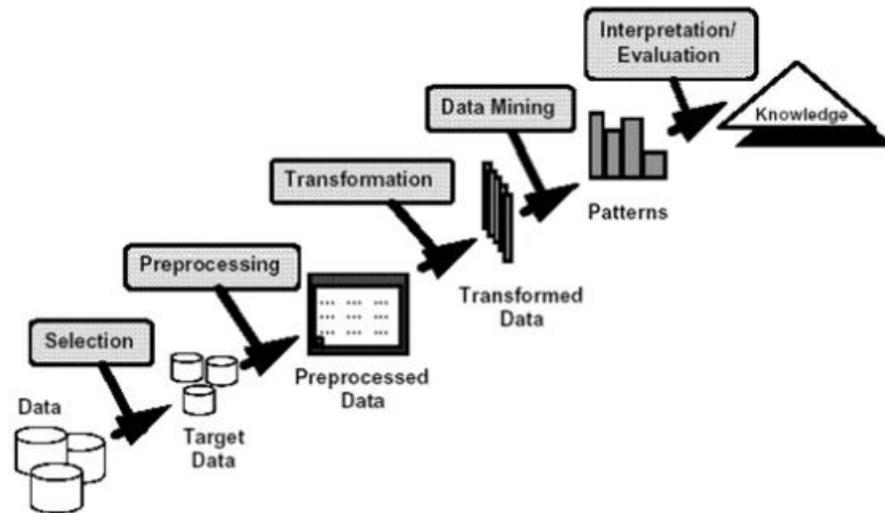
## **5.3 PROSES TAHAPAN DATA MINING**

Banyak teknik dan metode yang ada untuk melakukan berbagai jenis tugas data mining. Metode ini dikelompokkan dalam 3 paradigma utama data mining yaitu *Predictive Modeling*, *Discovery*, dan *Deviation Detection*. Data mining merupakan salah satu dari rangkaian *Knowledge Discovery In Database* (KDD). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah data [6].

Serangkaian proses tahapan data mining tersebut memiliki tahapan sebagai berikut :

1. Pembersih data (untuk membuang data yang tidak konsisten dan noise).
2. Integrasi data (penggabungan data dari beberapa sumber).
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining).
4. Aplikasi teknik data mining, proses ekstraksi pola dari data yang ada.
5. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan).
6. Presentasi pengetahuan (dengan teknik visualisasi)

Berikut gambar dari proses tahapan data mining :



**Gambar 2.1** Proses Tahapan Data Mining (*KDD*)

#### 5.4 ALGORITMA C4.5

Algoritma *C4.5* digunakan untuk membangun sebuah pohon keputusan yang mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar. Pohon keputusan adalah salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan [1].

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma *C4.5* yaitu :

1. Mempersiapkan *data training*, dapat diambil dari *data histori* yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menentukan akar dari pohon dengan menghitung nilai gain yang tertinggi dari masing-masing atribut atau berdasarkan nilai *index entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai *index entropy*, dengan rumus :

$$Entropy(i) = \sum_{j=1}^m f(i,j) \cdot \log_2 \frac{f(i,j)}{f(i)} \quad (1)$$

3. Hitung nilai gain dengan rumus :

$$Gain = - \sum_{i=1}^p \frac{n_i}{n} \cdot IE(i) \quad (2)$$

4. Untuk menghitung gain ratio perlu diketahui suatu term baru yang disebut Split Information dengan rumus :

$$SplitInformation = - \sum_{i=1}^c \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (3)$$

5. Selanjutnya menghitung gain ratio :

$$Gainratio(S,A) = \frac{Gain(S,A)}{SplitInformation(S,A)} \quad (4)$$

6. Ulangi langkah ke-2 hingga semua record terpartisi

Proses partisi pohon keputusan akan berhenti disaat :

- a. Semua tupel dalam record dalam simpul m mendapat kelas yang sama
- b. Tidak ada atribut dalam record yang dipartisi lagi
- c. Tidak ada record didalam cabang yang kosong.

## 5.5 ALGORITMA NAÏVE BAYES

Algoritma *Naive Bayes* adalah suatu klasifikasi kemungkinan sederhana yang dapat menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan. Suatu algoritma memanfaatkan teorema bayes dan memperkirakan seluruh atribut yang bebas dan saling lepas yang dapat diberikan oleh suatu nilai pada kelas variabel. *Naive Bayes* adalah klasifikasi dengan suatu metode kemungkinan dan perhitungan yang ditemukan oleh seseorang ilmuwan dari Inggris yaitu Thomas Bayes menghasilkan prediksi peluang yang akan datag berdasarkan suatu pengalaman sebelumnya [7]

Teorema Bayes memiliki bentuk umum sebagai berikut :

$$P(H \setminus X) = \frac{P(X \setminus H) P(H)}{P(H)} \dots\dots\dots (1)$$

Keterangan :

X = Kelas data yang belum diketahui.

H = Hipotesa data X adalah kelas spesifik.

$P(X|H)$  = Kemungkinan hipotesa H berdasarkan keadaan X (posteriori prob).

$P(H)$  = Kemungkinan Hipotesa H (prior prob).

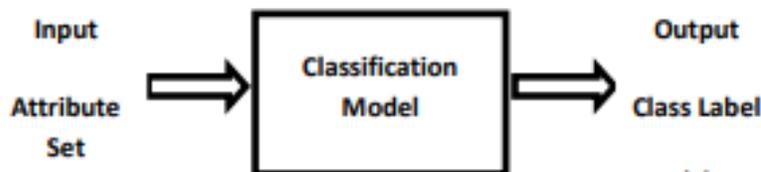
$P(X|H)$  = Kemungkinan X berdasarkan keadaan tersebut.

$P(X)$  = Kemungkinan dari X.

## 5.6 KLASIFIKASI

Klasifikasi merupakan proses pembelajaran suatu fungsi tujuan (target)  $f$  yang memetakan tiap himpunan atribut  $x$  ke satu dari label kelas  $y$  yang didefinisikan sebelumnya. Fungsi target disebut juga model klasifikasi.

Hal tersebut dapat digambarkan pada blok diagram sebagai berikut:



**Gambar 2.1 Blok Diagram Model Klasifikasi**

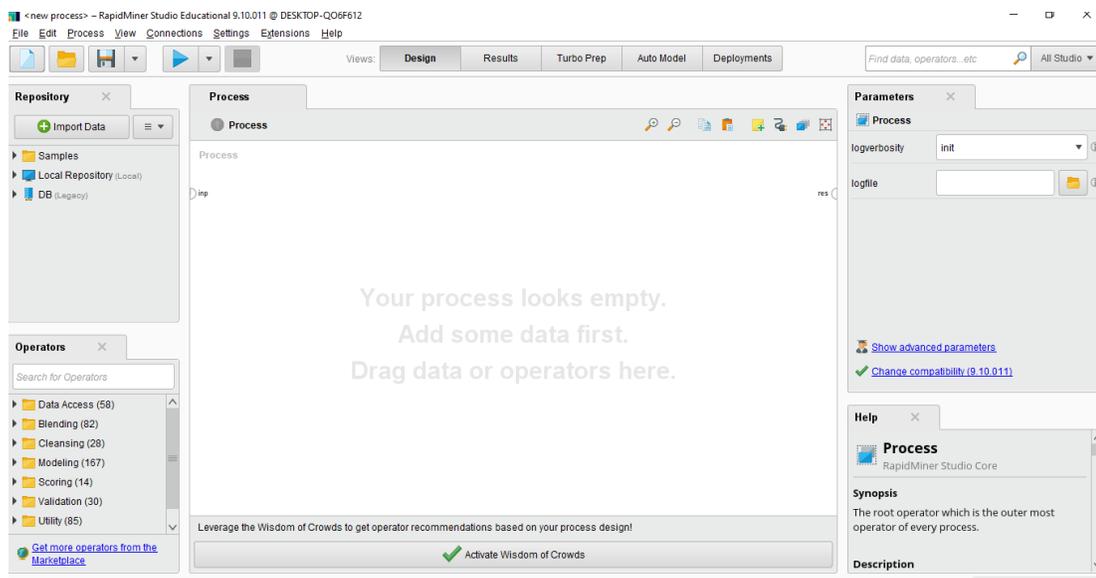
## 5.7 ALAT BANTU PENGUJIAN

### 5.7.1 RAPID MINER

*Rapid Miner* merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Ralf Klinkenberg dari rapid-i.com dengan tampilan GUI (*Graphical User Interface*) sehingga memudahkan pengguna dalam melakukan perangkat lunak ini. Perangkat lunak ini bersifat *open source* dan dibuat dengan menggunakan program Java di bawah lisensi *GNU Public Licence* dan *Rapid Miner* dapat dijalankan di sistem operasi apapun. Dengan

menggunakan *Rapid Miner*, tidak dibutuhkan kemampuan koding khusus, karena semua fasilitas sudah disediakan. *Rapid Miner* dikhususkan untuk penggunaan data mining. Model yang disediakan juga cukup banyak dan lengkap, seperti *Model Bayesin, Modelling, Tree Induction, Neural Network* dan lainnya.

Banyak metode yang disediakan *Rapid Miner* mulai dari Klasifikasi, Klustering asosiasi dan lainnya. Jika tidak ada model atau model algoritma yang tidak ada dalam *Rapid Miner*, pengguna boleh menambahkan modul lain, karena *Rapid Miner* bersifat *open source*, jadi siapapun dapat ikut mengembangkan perangkat lunak ini [8].

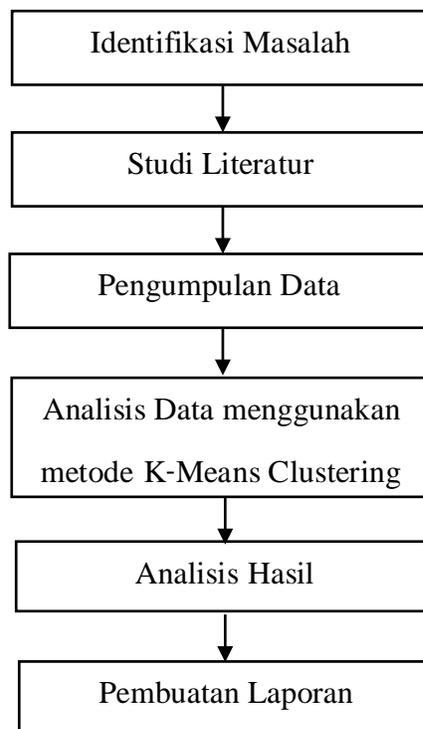


**Gambar 2.2** Tampilan Awal *Rapid Miner*

## 6. METODOLOGI PENELITIAN

### 6.1 KERANGKA KERJA PENELITIAN

Kerangka kerja penelitian merupakan proses atau tahapan yang akan dilakukan selama mengerjakan penelitian. Kerangka kerja penelitian dibuat dengan tujuan mempermudah kesuksesan hasil penelitian, dapat menyelesaikan penelitian tepat waktu dan penelitian dapat berjalan sesuai dengan yang diharapkan. Adapun kerangka kerja penelitian yang digunakan dapat dilihat pada gambar 3.1 berikut ini :



**Gambar 3.1 Kerangka Kerja Penelitian**

Berikut ini merupakan penjelasan dari setiap bagian yang tercantum pada kerangka kerja penelitian pada gambar 3.1 diatas :

#### a. Identifikasi Masalah

Pada tahap ini, penulis melakukan identifikasi dan merumuskan permasalahan pada penelitian, hal ini bertujuan agar mengetahui bagaimana memprediksi serta

mengukur tingkat akurasi penyakit gagal jantung. Data yang digunakan merupakan dataset yang di dapat dari laman <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction> Sehingga penulis mengetahui tahapan apa saja yang akan dilakukan dapat dalam penelitian ini.

#### **b. Studi Literatur**

Pada tahap ini, penulis melakukan pencarian terhadap landasan teori yang diperoleh dari berbagai sumber seperti buku, jurnal ilmiah dan juga dari berbagai referensi lainnya pencarian meliputi metode klasifikasi algoritma *C4.5* dan algoritma *Naïve Bayes*. Untuk melengkapi penelitian ini baik mengenai konsep dan teori sehingga memiliki acuan yang baik dan relevan.

#### **c. Pengumpulan Data**

Pada tahap ini, penulis melakukan pengumpulan data yang didapat melalui *repository dataset* yang berada pada *website* Kaggle.com. Kaggle merupakan tempat perkumpulan data *scientist* dimana tersedia banyak dataset yang dapat diambil dan tersedia banyak perlombaan yang berkaitan dengan *Machine Learning*. Pada data menjelaskan mengenai penyakit gagal jantung. Penyakit kardiovaskular (CVDs) adalah penyebab kematian nomor 1 secara global, mengambil sekitar 17,9 juta jiwa setiap tahun, yang menyumbang 31% dari semua kematian di seluruh dunia. Empat dari 5CVD kematian disebabkan oleh serangan jantung dan stroke, dan sepertiga dari kematian ini terjadi sebelum waktunya pada orang di bawah usia 70 tahun. Gagal jantung adalah kejadian umum yang disebabkan oleh CVD dan kumpulan data ini berisi 11 fitur yang dapat digunakan untuk memprediksi kemungkinan penyakit jantung. Orang dengan penyakit kardiovaskular atau yang berada pada risiko kardiovaskular tinggi (karena adanya satu atau lebih faktor risiko seperti hipertensi, diabetes, hiperlipidemia atau penyakit yang sudah ada) memerlukan deteksi dan manajemen dini di mana model pembelajaran mesin dapat sangat membantu.

Pada data ini terdapat 12 atribut dan 918 data, atribut yang terdiri dari :

1. Usia: usia pasien [tahun]

2. Jenis Kelamin: jenis kelamin pasien [M: Pria, F: Wanita]
3. Tipe Nyeri Dada: tipe nyeri dada [TA: Angina Khas, ATA: Angina Atipikal, NAP: Nyeri Non-Anginal, ASY: Tanpa Gejala]
4. RestingBP: tekanan darah istirahat [mm Hg]
5. Kolesterol: kolesterol serum [mm/dl]
6. PuasaBS: gula darah puasa [1: jika PuasaBS > 120 mg/dl, 0: sebaliknya]
7. EKG istirahat: hasil elektrokardiogram istirahat [Normal: Normal, ST: memiliki kelainan gelombang ST-T (inversi gelombang T dan/atau elevasi atau depresi ST > 0,05 mV), LVH: menunjukkan kemungkinan atau pasti hipertrofi ventrikel kiri menurut kriteria Estes]
8. MaxHR: detak jantung maksimum tercapai [Nilai numerik antara 60 dan 202]
9. Latihan Angina: angina yang diinduksi oleh olahraga [Y: Ya, N: Tidak]
10. Oldpeak: oldpeak = ST [Nilai numerik diukur dalam depresi]
11. ST\_Slope: kemiringan puncak latihan segmen ST [Up: upsloping, Flat: flat, Down: downsloping]
12. Penyakit Jantung: kelas keluaran [1: penyakit jantung, 0: Normal]

**d. Analisa Data**

Pada tahap ini penulis menganalisis hasil dari perhitungan metode Algoritma C4.5 dan Algoritma *Naïve Bayes* Klasifikasi, hasil analisis perhitungan penulis dengan hasil perhitungan menggunakan *tools Rapid Miner*.

**e. Pembuatan Laporan**

Pada tahap pembuatan laporan akhir penelitian, dilakukan berdasarkan kerangka yang telah dirancang yang terdiri dari Pendahuluan, Landasan Teori, Metodologi penelitian, Analisis dan Hasil, Hasil Analisis dan Visualisasi, dan Penutup serta lampiran bukti hasil penelitian.

## **6.2 METODE ANALISA DATA**

Metode yang digunakan adalah metode Algoritma *C4.5* dan Algoritma *Naïve Bayes* Klasifikasi yang mana *C4.5* dan *Naïve Bayes* merupakan salah satu metode data Klasifikasi. Klasifikasi merupakan pendekatan sistematis untuk pembuatan model klasifikasi dari sebuah data set input.

## **6.3 ALAT BANTU PENELITIAN**

Dalam melakukan penelitian, penulis juga memerlukan alat bantu. Alat bantu yang digunakan dalam melakukan analisis serta pengujian untuk mendapatkan hasil akurat agar sesuai dengan hasil yang diinginkan. Adapun alat bantu yang digunakan yaitu:

### **6.3.1 Perangkat Keras (*Hardware*)**

Perangkat keras yang penulis gunakan dalam penelitian ini adalah sebagai berikut :

1. Laptop : LENOVO
2. Prosesor : Intel(R) Core(TM) i3-5005U CPU @ 2.00GHz
3. RAM : 6 GB
4. Harddisk : 500 GB
5. Printer : Canon iP2770 series
6. Mouse : Moffi M3 Wireless

### **6.3.2 Perangkat Lunak (*Software*)**

Perangkat lunak yang penulis gunakan dalam penelitian ini adalah sebagai berikut :

1. Sistem Operasi Windows 10 64-bit
2. Microsoft Office 2016
3. *Tools Rapid Miner*
4. Google
5. Notepad++

### **6.3.3 Data yang digunakan dalam penelitian**

Berikut ini merupakan data yang akan digunakan dalam penelitian ini yang terdiri dari 918 data dan mempunyai 12 atribut.

Data tersebut diperoleh dari *website* Kaggle.com yang dapat dilihat pada laman <https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>

Dibawah ini merupakan tabel menggunakan beberapa data yang di dapat dari dataset. Dapat dilihat pada table berikut ini :

No	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	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
11	37	F	NAP	130	211	0	Normal	142	N	0	Up	0
12	58	M	ATA	136	164	0	ST	99	Y	2	Flat	1
13	39	M	ATA	120	204	0	Normal	145	N	0	Up	0
14	49	M	ASY	140	234	0	Normal	140	Y	1	Flat	1
15	42	F	NAP	115	211	0	ST	137	N	0	Up	0
16	54	F	ATA	120	273	0	Normal	150	N	1.5	Flat	0
17	38	M	ASY	110	196	0	Normal	166	N	0	Flat	1
18	43	F	ATA	120	201	0	Normal	165	N	0	Up	0
19	60	M	ASY	100	248	0	Normal	125	N	1	Flat	1
20	36	M	ATA	120	267	0	Normal	160	N	3	Flat	1
21	43	F	TA	100	223	0	Normal	142	N	0	Up	0
22	44	M	ATA	120	184	0	Normal	142	N	1	Flat	0
23	49	F	ATA	124	201	0	Normal	164	N	0	Up	0
24	44	M	ATA	150	288	0	Normal	150	Y	3	Flat	1
25	40	M	NAP	130	215	0	Normal	138	N	0	Up	0
26	36	M	NAP	130	209	0	Normal	178	N	0	Up	0
27	53	M	ASY	124	260	0	ST	112	Y	3	Flat	0
28	52	M	ATA	120	284	0	Normal	118	N	0	Up	0
29	53	F	ATA	113	468	0	Normal	127	N	0	Up	0
30	51	M	ATA	125	188	0	Normal	145	N	0	Up	0
31	53	M	NAP	145	518	0	Normal	130	N	0	Flat	1
32	56	M	NAP	130	167	0	Normal	114	N	0	Up	0

364	56	M	ASY	120	0	0	ST	148	N	0	Flat	1
365	56	M	NAP	120	0	0	Normal	97	N	0	Flat	0
366	64	F	ASY	200	0	0	Normal	140	Y	1	Flat	1
367	61	M	ASY	150	0	0	Normal	117	Y	2	Flat	1
368	68	M	ASY	135	0	0	ST	120	Y	0	Up	1
369	57	M	ASY	140	0	0	Normal	120	Y	2	Flat	1
370	63	M	ASY	150	0	0	Normal	86	Y	2	Flat	1
371	60	M	ASY	135	0	0	Normal	63	Y	0.5	Up	1
372	66	M	ASY	150	0	0	Normal	108	Y	2	Flat	1
373	63	M	ASY	185	0	0	Normal	98	Y	0	Up	1
374	59	M	ASY	135	0	0	Normal	115	Y	1	Flat	1
375	61	M	ASY	125	0	0	Normal	105	Y	0	Down	1
376	73	F	NAP	160	0	0	ST	121	N	0	Up	1
377	47	M	NAP	155	0	0	Normal	118	Y	1	Flat	1
378	65	M	ASY	160	0	1	ST	122	N	1.2	Flat	1
379	70	M	ASY	140	0	1	Normal	157	Y	2	Flat	1
380	50	M	ASY	120	0	0	ST	156	Y	0	Up	1
381	60	M	ASY	160	0	0	ST	99	Y	0.5	Flat	1
382	50	M	ASY	115	0	0	Normal	120	Y	0.5	Flat	1
383	43	M	ASY	115	0	0	Normal	145	Y	2	Flat	1
384	38	F	ASY	110	0	0	Normal	156	N	0	Flat	1
385	54	M	ASY	120	0	0	Normal	155	N	0	Flat	1
386	61	M	ASY	150	0	0	Normal	105	Y	0	Flat	1
387	42	M	ASY	145	0	0	Normal	99	Y	0	Flat	1
388	53	M	ASY	130	0	0	LVH	135	Y	1	Flat	1
389	55	M	ASY	140	0	0	Normal	83	N	0	Flat	1
390	61	M	ASY	160	0	1	ST	145	N	1	Flat	1
391	51	M	ASY	140	0	0	Normal	60	N	0	Flat	1
392	70	M	ASY	115	0	0	ST	92	Y	0	Flat	1
393	61	M	ASY	130	0	0	LVH	115	N	0	Flat	1
394	38	M	ASY	150	0	1	Normal	120	Y	0.7	Flat	1
395	57	M	ASY	160	0	1	Normal	98	Y	2	Flat	1
396	38	M	ASY	135	0	1	Normal	150	N	0	Flat	1

893	39	F	NAP	138	220	0	Normal	152	N	0	Flat	0
894	57	M	ATA	154	232	0	LVH	164	N	0	Up	1
895	58	F	ASY	130	197	0	Normal	131	N	0.6	Flat	0
896	57	M	ASY	110	335	0	Normal	143	Y	3	Flat	1
897	47	M	NAP	130	253	0	Normal	179	N	0	Up	0
898	55	F	ASY	128	205	0	ST	130	Y	2	Flat	1
899	35	M	ATA	122	192	0	Normal	174	N	0	Up	0
900	61	M	ASY	148	203	0	Normal	161	N	0	Up	1
901	58	M	ASY	114	318	0	ST	140	N	4.4	Down	1
902	58	F	ASY	170	225	1	LVH	146	Y	2.8	Flat	1
903	58	M	ATA	125	220	0	Normal	144	N	0.4	Flat	0
904	56	M	ATA	130	221	0	LVH	163	N	0	Up	0
905	56	M	ATA	120	240	0	Normal	169	N	0	Down	0
906	67	M	NAP	152	212	0	LVH	150	N	0.8	Flat	1
907	55	F	ATA	132	342	0	Normal	166	N	1.2	Up	0
908	44	M	ASY	120	169	0	Normal	144	Y	2.8	Down	1
909	63	M	ASY	140	187	0	LVH	144	Y	4	Up	1
910	63	F	ASY	124	197	0	Normal	136	Y	0	Flat	1
911	41	M	ATA	120	157	0	Normal	182	N	0	Up	0
912	59	M	ASY	164	176	1	LVH	90	N	1	Flat	1
913	57	F	ASY	140	241	0	Normal	123	Y	0.2	Flat	1
914	45	M	TA	110	264	0	Normal	132	N	1.2	Flat	1
915	68	M	ASY	144	193	1	Normal	141	N	3.4	Flat	1
916	57	M	ASY	130	131	0	Normal	115	Y	1.2	Flat	1
917	57	F	ATA	130	236	0	LVH	174	N	0	Flat	1
918	38	M	NAP	138	175	0	Normal	173	N	0	Up	0

## 7. JADWAL PENELITIAN

Berikut jadwal penelitian agar terhindar dari ketidak teraturan dalam proses pengerjaan penelitian ini. Disusunlah jadwal penelitian yang direncanakan akan dilakukan selama 4 (empat) bulan, yaitu dari bulan **September 2022** sampai dengan bulan **Desember 2022** dengan jadwal sebagai berikut :

KEGIATAN	BULAN																			
	September				Oktober				November				Desember				Januari			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Identifikasi Masalah	■	■																		
Studi Literatur			■	■																
Pengumpulan Data dan Informasi					■	■	■	■												
Analisis Data Mining						■	■	■	■	■	■	■								
Penerapan Data Mining									■	■	■	■	■	■						
Penyusunan Data									■	■	■	■	■	■	■					
Pembuatan Laporan					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

## DAFTAR PUSTAKA

- [1] 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.

- [2] D. A. Langga and Dkk, “Perbandingan Algoritma Naive Bayes Dengan Algoritma K-Nearest Neighbor Untuk Prediksi Penyakit Jantung,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [3] M. O. Deswita, “Pemeriksaan Penunjang Heart Failure,” *Universitas Padjadjaran*, vol. 4, no. 2, pp. 25–34, 2016.
- [4] I. A. Darmawan, M. F. Randy, I. Yuniarto, M. M. Mutoffar, and M. T. P. Salis, “Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Pola Golongan Penyandang Masalah Kesejahteraan Sosial,” *Sebatik*, vol. 26, no. 1, pp. 223–230, 2022, doi: 10.46984/sebatik.v26i1.1622.
- [5] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database ( KDD ) . Jurnal Edik Informatika,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [6] “Proses proses tahapan data mining - Dokterskripsi.” [Online]. Available: <https://dosbing.id/2022/02/27/proses-tahapan-data-mining/>
- [7] R. Rachman and R. N. Handayani, “Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,” *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.
- [8] A. Supriyanto, “Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa Menggunakan Metode C4. 5 (Studi Kasus PTIK FKIP UNS),” *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2020.