

**PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK  
PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN  
*LIBRARY TENSORFLOW***

**SKRIPSI**

**Diajukan Oleh**

**MUHAMMAD RISYAD RAHMADI**

**NIM. 200212062**

**Bidang Peminatan : Rekayasa Perangkat Lunak (RPL)**

**Mahasiswa Fakultas Tarbiyah dan keguruan**

**Program Studi Pendidikan Teknologi Informasi**



**UNIVERSITAS ISLAM NEGERI AR-RANIRY**

**FAKULTAS TARBIYAH DAN KEGURUAN**

**PROGRAM STUDI PENDIDIKAN TEKNOLOGI INFORMASI**

**2024 M/ 1445 H**

# SKRIPSI

## **PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN *LIBRARY TENSORFLOW***

Oleh :

**Muhammad Risyad Rahmadi**

**Mahasiswa Fakultas Tarbiyah dan Keguruan**

**Program Studi Pendidikan Teknologi Informasi**

**NIM. 200212062**

**Bidang Peminatan : Rekayasa Perangkat Lunak (RPL)**

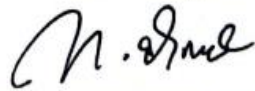


Disetujui Oleh :

Pembimbing 1

  
(Nazaruddin Ahmad, M.T.)  
NIP/NIDN : 0105068202

Pembimbing 2

  
(Nazaruddin Ahmad, M.T.)  
NIP/NIDN : 0105068202

# PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN *LIBRARY TENSORFLOW*

## SKRIPSI

Telah diuji Oleh Panitia Ujian Munaqasyah Skripsi Fakultas Tarbiyah dan keguruan UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus serta diterima sebagai satu beban studi Program Sarjana (S-1) dalam Pendidikan Teknologi Informasi

Pada :

Kamis, 08 Agustus 2024  
03 Shafar 1446 H

Darussalam – Banda Aceh  
Panitian Ujian Munaqasyah Skripsi

Ketua

(Nazaruddin Ahmad, M.T)  
NIP.NIDN. 0105068202

Sekretaris

(Nazaruddin Ahmad, M.T)  
NIP.NIDN. 0105068202

Penguji 1

(Baihaqi, M.T)  
NIP.NIDN.198402242019031004

Penguji 2

(Aulia Syarif Aziz, S.Kom., M.Sc.)  
NIP.NIDN.199305212022031001

Mengetahui,

Dekan Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry  
Darussalam, Banda Aceh



(Prof. Saffrudin Muluk, S.Ag., M.A., M.Ed., Ph.D.)  
NIP. (197301021997031003)

(Prof. Sarfrul Muluk, S.Ag., M.A., M.Ed., Ph.D.)

NIP.197301021997031003

### LEMBAR PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan dibawah ini :


Nama : Muhammad Risyad Rahmadi  
NIM : 200212062  
Program Studi : Pendidikan Teknologi Informasi  
Fakultas : Tarbiyah dan Keguruan  
Judul Skripsi : Penerapan Algoritma *Naive Bayes* Untuk Prediksi Penyakit Jantung menggunakan *Library Tensorflow*


Dengan ini menyatakan bahwa dalam penulisan skripsi ini, saya :

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan
2. Tidak melakukan plagiat terhadap naskah karya orang lain
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya
4. Tidak memanipulasi dan memalsukan data
5. Mengerjakan sendiri karya ini dan mampu bertanggung jawab atas karya ini

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat di pertanggung jawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh.

Demikian pernyataan ini saya buat dengan sesungguhnya

Banda Aceh , 08 .Agustus 2024  
Yang menyatakan,  
  
Muhammad Risyad Rahmadi  
NIM. 200212062



## ABSTRAK

Nama : Muhammad Risyad Rahmadi  
NIM : 200212062  
Program Studi : Pendidikan Teknologi Infotmasi  
Judul : Penerapan Algoritma *Naive Bayes* Untuk Prediksi Penyakit Jantung Menggunakan *Library* Tensorflow  
Bidang Peminatan : Rekayasa Perangkat Lunak (RPL)  
Jumlah Halaman : 61 Halaman  
Pembimbing Utama : Nazaruddin Ahmad, M.T.  
Kata kunci : Algoritma *Naive Bayes*, Python, *Machine Learning* Penyakit Jantung.

Ilmu pengetahuan saat ini menjangkau macam-macam bidang salah satu pemanfaatan ilmu pengetahuan, yang dipakai untuk pengolahan data dibidang kesehatan. Salah satu contohnya Ilmu pengetahuan di bidang kesehatan pada data pasien penyakit jantung , namun sebagian besar data pasien tersebut tidak diolah dengan bijak sehingga mendapatkan informasi yang kurang efektif untuk pengambilan keputusan kesehatan pasien yang terkena penyakit jantung maupun tidak. Dalam pengambilan keputusan tersebut dengan dasar data atau informasi yang akurat akan menghasilkan sebuah keputusan atau prediksi terkait dengan penyakit maupun kesehatan pasien tersebut menjadi lebih tepat dengan dasar data yang sudah ada. Dengan permasalahan yang ada maka dibuatkan sebuah sistem prediksi yang mampu memprediksi penyakit jantung tersebut dengan berlandaskan data yang sudah ada dengan menggunakan algoritma *Naive Bayes* menggunakan bahasa pemrograman python agar memudahkan dalam memprediksi atau pengambilan sebuah keputusan yang tepat berlandaskan data yang sudah ada terjadi sebelumnya. pada penelitian ini penulis menggunakan metode penelitian kuantitatif dengan pendekatan kuantitatif deskriptif dan hasil akhir yang didapatkan adalah sebuah sistem berbasis codingan yang mampu memprediksi data yang terkena penyakit jantung tersebut dengan tingkat akurasi yang tinggi di angka 82%, 83% dan 84% dari dataset berjumlah 1024 data penyakit jantung.

## KATA PENGANTAR

Puji syukur diucapkan kehadiran Allah SWT atas segala rahmatNya sehingga Proposal Skripsi ini dapat tersusun sampai dengan selesai. Tidak lupa kami mengucapkan terimakasih terhadap bantuan dari pihak yang telah berkontribusi dengan memberikan sumbangan baik pikiran maupun materinya. Penulis mengucapkan terima kasih kepada:

1. Kedua orang tua, Bapak dan Ibu yang telah memberikan segalanya selama menjalani Pendidikan.
2. Bapak Dekan Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh.
3. Ibu Mira Maisura, M.Sc selaku Ketua Program Studi Pendidikan Teknologi Informasi atas kesempatan dan bantuan yang diberikan kepada penulis dalam melakukan penelitian dan memperoleh informasi yang diperlukan selama penulisan proposal penelitian ini.
4. Bapak Nazaruddin Ahmad, M.T sebagai Dosen Pembimbing Skripsi yang telah memberikan arahan dan semangat dalam penyusunan Skripsi
5. Bapak/Ibu Dosen program studi Pendidikan Teknologi Informasi yang telah mendidik dan memberikan bimbingan selama masa perkuliahan.

Meskipun telah berusaha menyelesaikan skripsi ini sebaik mungkin, penulis menyadari bahwa skripsi ini masih ada kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun terima kasih .Rahmat dan hidayah-Nya kepada kita semua. Aamiin ya rabbal 'alamin.

Banda Aceh, 08 Agustus 2024

Penulis



Muhammad Risyad Rahmadi

## DAFTAR ISI

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>LEMBAR PENGESAHAN PEMBIMBING</b> .....	<b>ii</b>
<b>LEMBAR PENGESAHAN SIDANG</b> .....	<b>iii</b>
<b>LEMBAR PERNYATAAN KEASLIAN</b> .....	<b>iv</b>
<b>ABSTRAK</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>vi</b>
<b>DAFTAR ISI</b> .....	<b>vii</b>
<b>DAFTAR GAMBAR</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>x</b>
<b>DAFTAR PERSAMAAN</b> .....	<b>xi</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian .....	3
1.4 Batasan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Penelitian Terdahulu .....	4
<b>BAB II LANDASAN TEOROTIS</b> .....	<b>7</b>
2.1 Penyakit Jantung .....	7
2.2 Data Mining .....	8
2.2.1 Pengelompokan Data Mining.....	9
2.2.2 Tahapan Data Mining.....	11
2.3 Klaisfikasi .....	13
2.4 Naïve Bayes .....	14
2.5 Probabilitas.....	15
2.6 Confusion Matrix .....	15
2.7 Tensorflow .....	18
<b>BAB III METODE PENELITIAN</b> .....	<b>19</b>
3.1 Pendekatan Penelitian .....	19
3.2 Subjek Penelitian serta Sumber Data .....	19
3.2.1 Subjek Penelitian.....	19
3.2.2 Sumber Data.....	19
3.3 Teknik Pengumpulan Data.....	19
3.3.1 Website <a href="http://www.Kaggle.com">www.Kaggle.com</a> .....	19
3.4 Teknik Analisis Data.....	20
3.5 Rancangan Penelitian .....	20

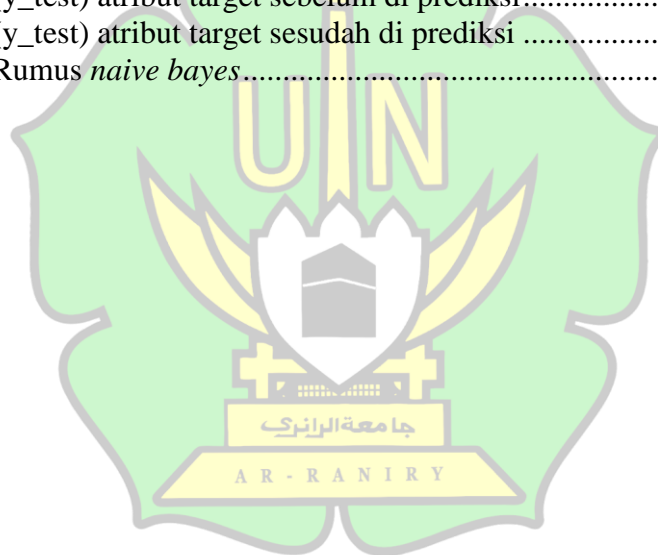
<b>BAB IV .....</b>	<b>26</b>
<b>HASIL DAN PEMBAHASAN .....</b>	<b>26</b>
4.1. Hasil Penelitian .....	26
4.1.1. Persiapan Data .....	26
4.2. Pemilihan Atribut .....	30
4.2.1. Pemilihan Atribut .....	30
a). Data Cleaning .....	31
b). Data Split.....	31
 <b>BAB V .....</b>	 <b>44</b>
<b>KESIMPULAN DAN SARAN .....</b>	<b>44</b>
5.1 Kesimpulan .....	44
5.2 Saran.....	44
 <b>DAFTAR PUSTAKA .....</b>	 <b>46</b>
<b>RIWAYAT HIDUP PENULIS .....</b>	<b>76</b>





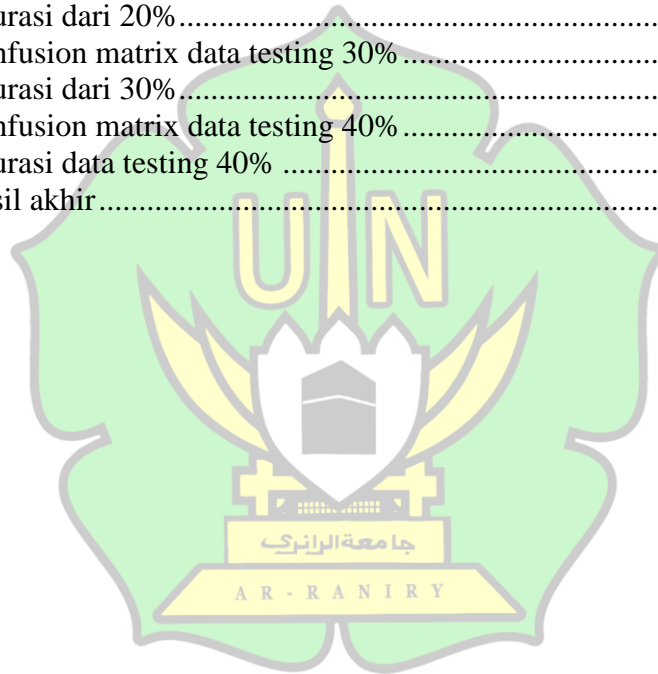
## DAFTAR GAMBAR

Gambar 1. 1 Tahapan Data Mining.....	11
Gambar 3. 1 Alur Analisi Data. ....	<b>20</b>
Gambar 3. 2 Alur Penelitian.....	20
Gambar 3. 3 Tahapan pengolahan data.....	24
Gambar 4. 1 Import library ( <i>sumber pribadi</i> ) .....	<b>26</b>
Gambar 4. 2 Code membaca data .....	27
Gambar 4. 3 Grafik jumlah yang terkenapenyakit jantung. ....	28
Gambar 4. 4 Gambar grafik dari jenis kelamin.....	29
Gambar 4. 5 Grafik dari segi umur pasien .....	29
Gambar 4. 6 ( <i>y_test</i> ) atribut target sebelum di prediksi.....	34
Gambar 4. 7 ( <i>y_test</i> ) aatribut target yang sudah diprediksi .....	34
Gambar 4. 8 ( <i>y_test</i> ) atribut target sebelum di prediksi.....	35
Gambar 4. 9 ( <i>y_test</i> ) atribut target sesudah di prediksi .....	36
Gambar 4. 10 ( <i>y_test</i> ) atribut target sebelum di prediksi.....	37
Gambar 4. 11 ( <i>y_test</i> ) atribut target sesudah di prediksi .....	38
Gambar 4. 12.Rumus <i>naive bayes</i> .....	39



## DAFTAR TABEL

Tabel 1.1 Penelitian Terdahulu .....	4
Tabel 2.1 Tabel Confusion Matrix.....	16
Tabel 4 1.Dataset.....	27
Tabel 4 2. Atribut .....	30
Tabel 4 3. Cleaning data (pembersihan data).....	31
Tabel 4 4. Data testing 20% .....	32
Tabel 4 5. Variable dependen (y_test) .....	32
Tabel 4 6. Data Testing 30% .....	34
Tabel 4 7. Variable dependen (y_test) .....	35
Tabel 4 8. Data Testing 40% .....	36
Tabel 4 9. Variable dependen (y_test) .....	37
Tabel 4 10. Cofusion matrix data testing 20% .....	39
Tabel 4 11. akurasi dari 20%.....	40
Tabel 4 12. confusion matrix data testing 30% .....	40
Tabel 4 13. akurasi dari 30%.....	41
Tabel 4 14. confusion matrix data testing 40% .....	42
Tabel 4 15. akurasi data testing 40% .....	43
Tabel 4 16. hasil akhir.....	43



## DAFTAR PERSAMAAN

(Persamaan 1 Teorema Naive Bayes.) .....	14
(Persamaan 2 Rumus Probabilitas.....)	15
(Persamaan 3 Rumus Akurasi).....)	17
(Persamaan 4 Rumus Recall). .....	17
(Persamaan 5 Rumus Precision). .....	17
(Persamaan 6 Rumus F1-Score).....)	17



## DAFTAR LAMPIRAN

Lampiran 1 Surat Keterangan Skripsi.....	49
Lampiran 2 Form Bimbingan.....	50
Lampiran 3 Data Testing.....	51



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Ilmu pengetahuan saat ini menjangkau macam-macam bidang salah satunya dipakai untuk pengolahan data di bidang kesehatan. Salah satu contohnya di bidang kesehatan pada data pasien penyakit jantung, namun sebagian besar data pasien tidak diolah dengan bijak sehingga mendapatkan informasi yang kurang efektif untuk pengambilan keputusan kesehatan pasien yang terkena penyakit jantung maupun tidak.

Pada umumnya prediksi yang dilakukan oleh dokter untuk membuat keputusan terhadap pasien sering dilakukan dengan intuisi bukan dengan terhadap data yang sudah terjadi pada sebelumnya maka akan terjadi ketidaktepatan terhadap pengambilan keputusan.

Jantung adalah organ penting pada tubuh manusia karena fungsi jantung adalah memompa darah ke seluruh tubuh melalui pembuluh darah. Penyakit Jantung merupakan sebutan umum yang beredar dikalangan masyarakat yang digunakan untuk menggambarkan gangguan terhadap fungsi kerja jantung. Terdapat juga beberapa serangkaian gangguan yang mempengaruhi fungsi jantung tersebut yaitu penyakit jantung koroner, penyakit jantung bawaan bahkan gagal jantung [1].

Dalam hal ini akan membuat sebuah keputusan tidak tepat dalam pengambilan keputusan maka dari itu dalam perkembangan ilmu pengetahuan sekarang ini, kita memiliki kesempatan bisa untuk memprediksi pasien yang

terkena penyakit jantung dari tekanan darah, jenis sakit dada, kolesterol dan sebagainya. Maka yang diperlukan adalah sebuah ilmu pengetahuan untuk membuat sistem yang memprediksi apakah orang tersebut terkena penyakit jantung atau tidak dengan dasar data yang sudah ada terlebih dahulu [1].

Adapun permasalahan yang diangkat oleh penulis adalah untuk membandingkan hasil dari penelitian sebelumnya dengan algoritma *Naïve Bayes* dan data yang dipakai berjumlah 1024 dataset [1].

Penulis menawarkan solusi dengan permasalahan yang timbul akan membuat sebuah sistem berbasis codingan untuk memprediksi pasien terkena penyakit jantung atau tidak berdasarkan dengan data yang sudah ada menggunakan algoritma *Naive Bayes*. Algoritma *Naive Bayes* juga memiliki kelebihan diantaranya mudah diimplementasi dan memberikan hasil yang baik untuk data yang berjumlah besar [2].

Dataset penyakit jantung yang di diambil dari sebuah website yang bernama *www.kaggle.com* dengan jumlah berkisar 1024 dataset penyakit jantung dan mencari akurasi yang didapatkan apakah lebih bagus/tinggi dari algoritma yang dipakai pada penelitian sebelumnya yang menjadi topik utama dari penelitian.

## **1.2 Rumusan Masalah**

Rumusan masalah pada penelitian ini ada beberapa yaitu :

- 1) Bagaimana cara penerapan algoritma *Naïve Bayes* pada prediksi penyakit jantung ?
- 2) Bagaimana melihat hasil penerapan algoritma *Naïve Bayes* tersebut dengan data *training* 1024 dataset dengan data *testing* sebesar 20%, 30% dan 40% ?

### 1.3 Tujuan Penelitian

Tujuan penelitian yang akan didapatkan pada penelitian ini ada beberapa yaitu :

- 1) Untuk Mengetahui penerapan algoritma *Naïve bayes*.
- 2) Untuk membandingkan hasil dari akurasi yang didapatkan pada algoritma *Decision tree* dengan Algoritma *Naive Bayes* yang dipakai pada penelitian ini.

### 1.4 Batasan Penelitian

Batasan pada penelitian yang akan dilakukan ada beberapa yaitu :

- 1) Memakai Perhitungan *Naïve Bayes*.
- 2) Menggunakan data penyakit jantung sebagaia dataset yang di dapatkan dari website [www.kaggle.com](http://www.kaggle.com) yang bernama data *Heart Diaese*.
- 3) Data *Testing* digunakan sebesar 20%, 30% dan 40% dari data *Training*.
- 4) Atribut yang dipakai adalah *sex, age, cp, chol, fbs, trestbps,,resteg, thalach, slope, exang,,ca, thal oldpeak, condition*.
- 5) Alat bantu penelitian Bahasa pemrograman *Python* dan *library tensorflow*.

### 1.5 Manfaat Penelitian

Manfaat pada Penelitian yang dilakukan bagi penulis dan pembaca yaitu :

- 1) Meningkatkan pengalaman maupun pengetahuan pada data *mining* dan keinginan dari penulis menjadi seorang *Data Analyst*.

Manfaat Penelitian bagi Pembaca :

- 1) Mengetahui kinerja dari Algoritma *Naïve bayes* dan bisa dilanjutkan untuk ide penelitian baru.

## 1.6 Penelitian Terdahulu

Tabel 1. 1 Penelitian Terdahulu

No	Judul Penelitian Sebelumnya	Tahun Penelitian	Entitas Penelitian	Hasil Penelitian
1	Perbandingan Algoritma Klasifikasi dalam Pengklasifikasian Data Penyakit Jantung Koroner.	2019	Data penyakit jantung koroner, Algoritma , <i>K-Nearest Neighbor</i> , <i>DecisionTree</i> dan <i>Random Forest</i> , <i>Naive Bayes</i> metode metode <i>k-fold cross validation</i>	Berdasarkan Penelitian tersebut bahwasanya penerapan algoritma tersebut dapat direalisasikan, Algoritma <i>Random Forest</i> yang menghasilkan akurasi sekitar kurang lebih 85,668%. Untuk kelas '1', recall mencapai 89%, sedangkan untuk kelas '0'(nol), <i>recallnya</i> adalah 83,6%. Selain itu, <i>presisi</i> untuk kelas '1'(satu) mencapai 85%. Kesimpulannya bahwa Algoritma <i>Random Forest</i> terbukti menjadi pilihan terbaik dalam mengklasifikasi kasus penyakit jantung koroner jika dibandingkan dengan algoritma <i>Naïve Bayes</i> , <i>K-Nearest Neighbor</i> , dan <i>Decision Tree</i> .
2	Analisis Perbandingan Algoritma Klasifikasi untuk Penyakit Jantung.	2023	Data penyakit jantung, algoritma <i>Naive Bayes</i> , <i>Decision Tree</i> , <i>Random Forest Classifier</i> serta kinerja algoritma	Hasil Pengujian Berhasil, dari ketiga algoritma tersebut. Algoritma <i>Decision tree</i> mendapatkan akurasi sebesar 0,844 menggunakan random search dan menggunakan grid sebesar 0,84. Algoritma <i>Naive Bayes</i> mendapatkan akurasi sebesar 0,85 dengan kedua metode tersebut. Dan <i>Random Forest</i> mendapatkan akurasi sebesar 0,852 menggunakan random search dan 0.868 menggunakan grid.
3	Analisis Penyakit Jantung Koroner Menggunakan <i>Decision Tree</i> .	2021	Data penyakit jantung koroner, Algoritma <i>Decision Tree</i> , <i>Confusion Matrix</i>	Berdasarkan penelitian dan uji coba pada dataset yang tersedia, model pohon keputusan <i>CART</i> mampu mengklasifikasikan penyakit jantung koroner dengan



No	Judul Penelitian Sebelumnya	Tahun Penelitian	Entitas Penelitian	Hasil Penelitian
				akurasi sekitar 80%. Proses ini melibatkan beberapa tahapan, termasuk tahap post-pruning.
4	Klasifikasi Penyakit Jantung Menggunakan <i>Decision Tree</i> dan <i>Random Forest</i> .	2023	Data penyakit jantung, <i>K-fold Cross Validation</i> , <i>Algoritma Decision Tree</i> , <i>Random Forest</i> ,	Dampak yang dihasilkan analisis menunjukkan bahwa klasifikasi penyakit jantung menggunakan <i>Random Forest</i> lebih baik daripada <i>Decision Tree</i> . Klasifikasi dengan <i>Random Forest</i> memberikan hasil yang lebih akurat, dengan akurasi sekitar 81.82%. Selain itu, presisi mencapai 87.04%, recall sekitar 79.82%, dan F1 score sekitar 83.13%.
5	Klasifikasi <i>Algoritma Naive Bayes</i> Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM.	2021	Data sewa UMKM, aplikasi <i>Rapidminer Data Mining</i> dan <i>Naive Bayes</i> .	Dampak dari ini adalah Hasil pengolahan data menggunakan aplikasi <i>Rapidminer</i> dengan <i>algoritma Naive Bayes</i> menghasilkan macam-macam informasi berikut: Akurasi sebesar 81.81%, Presisi sekitar 66.66%, Recall mencapai 100%, AUC (Area Under the Curve) sebesar 0.800%. Pengolahan ini dibuat untuk mengidentifikasi jumlah status pembayaran yang <i>outstanding</i> dari beberapa kriteria. Ini bermanfaat membantu manajemen untuk membuat langkah-langkah guna mengurangi kasus seperti penyewa yang selalu telat dalam pembayaran sewa teras tersebut.
6	Aplikasi <i>Metode Naive Bayes</i> dalam Prediksi Risiko Penyakit Jantung.	2017	Data penyakit jantung, <i>Algoritma Naive Bayes</i> , <i>Classification</i> , <i>Data Mining</i>	Dampak dari penelitian ini adalah pada <i>Metode Naive Bayes</i> yang berfungsi untuk memprediksi risiko terkena penyakit jantung. Pengaplikasian metode <i>Naive Bayes</i> dalam memprediksi risiko penyakit jantung menggunakan program <i>Delphi 7 Enterprise</i> dengan nama aplikasi HDP v1.0 telah berhasil dilakukan. Hasil prediksi akurat dalam pengelompokan risiko orang yang terkena penyakit jantung berdasarkan data dari RSUD

No	Judul Penelitian Sebelumnya	Tahun Penelitian	Entitas Penelitian	Hasil Penelitian
				<p>AWS tahun 2016 telah terbukti. Pada percobaan 1 dengan 25 <i>data testing</i>, tingkat akurasi mencapai 80%, sedangkan pada percobaan 2 dengan 50 <i>data testing</i>, tingkat akurasi mencapai 78%.</p>



## **BAB II**

### **LANDASAN TEOROTIS**

#### **2.1 Penyakit Jantung**

Salah satu sistem paling inti pada manusia sendiri adalah sistem peredaran darah itu sendiri . Adapun 2 fungsi utama pada sistem ini diantaranya berfungsi sebagai membagi oksigen maupun nutrisi kepada organ keseluruhan pada tubuh manusia dan berfungsi juga mengangkat hasil sisa dari metabolisme ditubuh. Dapat disimpulkan bahwa organ tubuh yang sangat penting akan kegunaanya pada sistem peredaran darah tubuh manusia adalah jantung [1].

Penyakit Jantung juga merupakan penyakit yang disebabkan oleh penyumbatan pada arteri koroner oleh tumpukan plak, polutan atau zat-zat kimia yang berasal dari lingkungan yang biasanya masuk ke tubuh melalui makanan, minuman. Hal ini dapat mengakibatkan adanya kemungkinan akan terjadi penggumpalan darah pada bagian arteri yang menyempit, dengan begitu tidak ada lagi darah yang akan mengalir karena aliran arteri dihalang oleh gumpalan darah tersebut yang sudah menjadi keras [3].

Jantung adalah adalah suatu organ tubuh dari manusia yang berperan dalam sistem peredaran darah. Ketika otot jantung tidak memiliki aliran darah yang cukup, maka sindrom koroner akut atau serangan jantung dapat mengganggu kerja jantung dengan serius, sehingga fungsi jantung untuk mengedarkan darah menuju seluruh tubuh menjadi terganggu [4].

## 2.2 Data Mining

Data Mining atau penambangan data merupakan proses pencarian pola yang menarik dan tersembunyi dari kumpulan data yang berukuran besar. Data tersebut tersimpan di basis data, Gudang data, atau tempat penyimpanan data lainnya. Penambangan data merupakan salah satu bagian dari kecerdasan buatan [5].

Data mining juga merupakan proses menemukan dan menganalisis data dalam jumlah besar dengan tujuan menemukan pola atau informasi yang menarik dari sejumlah besar data yang tersimpan dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses knowledge discovery in database (KDD) secara keseluruhan. Hasil dari proses data mining dapat digunakan sebagai evaluasi pengambilan keputusan dimasa depan [6]

Pengertian lain dari *data mining* itu sendiri adalah satu proses didalamnya memakai beberapa cara yang pertama statistik, *Machine Learning* dan sebagainya yang memiliki tujuan untuk menyaring dan mengetahui atau mendapatk informasi sangat berguna serta juga pembaharuan baru atau ilmu baru yang terikat di berbagai *database* (penyimpanan data) sangat besar. Pengertian lainnya merupakan sebuah operasi yang melibatkan teknik ataupun metode tertentu dengan tujuan mencari pola dan bentuk yang berbeda didalam sebuah data yang terpilih [7].

Disisi lain *Data Mining* juga memiliki makna sebagai salah satu pekerjaan yang mengenai pengumpulan maupun pemakaian pada data historis yang bertujuan mendapatkan suatu pola maupun kaitannya dalam dataset yang memiliki ukuran

begitu besar. Penjelasan diatas dapat disimpulkan bahwasannya *data mining* merupakan suatu teknik pengumpulan maupun pemakaian pada data yang menggunakan ilmu statistik dan *machine learning* pada hal-hal mengenai mengetahui informasi-informasi yang berguna [8].

### 2.2.1 Pengelompokan *Data Mining*

Pendapat dari Gunadi et al (2016), beberapa pengelompokannya sebagai berikut :

#### a) Deskripsi

Gambaran atau biasa disebut gambaran pola dan keseringang memberikan pemahaman-pemahaman suatu pola atau kecendrungan yang terdapat pada dalam data tersebut.

#### b) Klasifikasi

Klasifikasi atau dikenal sebagai penggolongan adalah satu metode yang mampu menjabarkan proses untuk penemuan model bertujuan menggambarkan serta pembedaan kelas pada data yang bertujuan memprediksi kelas dari objek ynag data kelasnya samar-samar atau masih diragukan. Metode ini terbagi menjadi 4 komponen adalah *Predictor, class, pelatihan data* serta mencoba dataset.

#### c) Estimasi

Estimasi pengertian atau penjelasannya mendekati dengan *Clasification* atau penggolongan diatas ,akan terkecuali jika *variable* tujuan estimasi yang telah diterapkan akan cenderung ke 'angka' daripada ke kategori.

Gambaran dibentuk dengan cara melalui catatan lengkap yang telah disediakan *value* (nilai) dari *variabel* tujuan acuan *value* prediksi.

d) Prediksi

Penjelasan Prediksi juga hampir sama dengan *Clasification* dan estimasi, terkecuali bahwasannya dalam prediksi nilai hasil akan ada di masa yang mendatang. Misalkan prediksi dalam kesehatan seperti kenaikan angka penyakit jantung 4 tahun mendatang. Secara singkatnya adalah memperkirakan atau meramal pada masa yang akan dilalu berlandaskan data yang telah didapatkan terdahulu.

e) Pengklusteran (kelompok)

Clustering atau biasa disebut pengelompokan data atau biasa disebut record , pemantauan serta membuat kelas dari masing-masing objek yang mempunyai kesamaan. *Cluster* (kelompok) merupakan himpunan-himpunan data yang mempunyai kesamaan antara satu sama lain serta juga mempunyai ketidaksamaan antara data-adata atau macam-macam record didalam kelompok lain (*cluster*). Pengelompokkan ini berlainan dengan pengklasifikasian (*clasification*) karena tidak mempunyai *variabel* tujuan di kelompok (*cluster*). Pengelompokkan tidak akan melakukan *clasification*, nilai(angka), bahkan mengetahui hasil yang akan datang(prediksi) sebuah nilai dari *variable* tujuan.

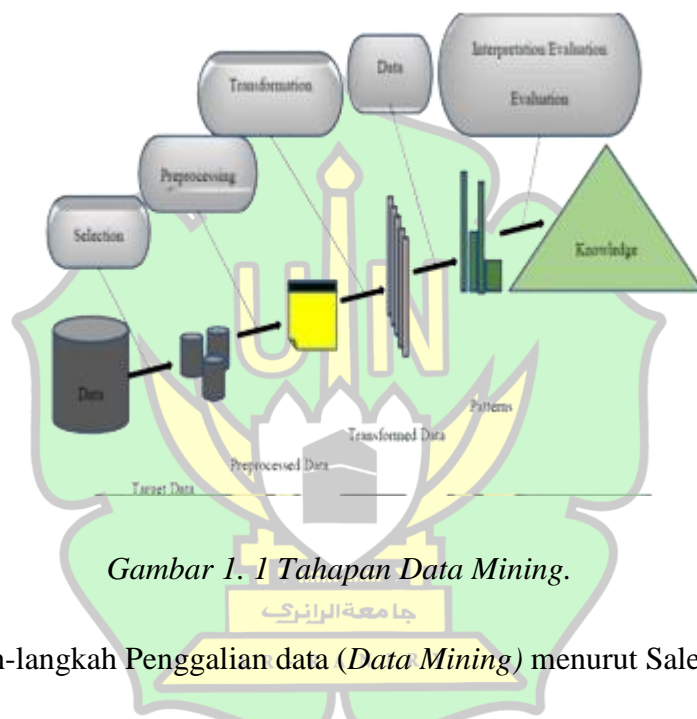
f) Asosiasi

Kumpulan atau disebut dengan asosiasi tujuanya adalah menemukan identitas yang keluar di satu waktu yang sama. Pada lingkungan

perbisnisan biasa disebut analisa pemberlanjaan. Contoh singkat kumpulan (asosisasi) pada bisnis adalah mendapatkan suatu barang yang ada pada toko penjual tersebut di beli dengan waktu bersamaan dan begitu juga tidak di beli pada waktu yang bersamaan [9].

### 2.2.2 Tahapan *Data Mining*

Berikut adalah gambar alurnya :



*Gambar 1. 1 Tahapan Data Mining.*

langkah-langkah Penggalan data (*Data Mining*) menurut Saleh (2015) :

1. Penyaringan atau pembersihan data pengertiannya merupakan Pembersihan atau penyaringan data dilakukan dengan menghilangkan data-data yang rusak atau tidak sesuai dan memperbarui data agar berkualitas.
2. Penyatuan data atau *Data Integration*  
Merupakan penggabungan data contohnya data yang berbeda asal-usulnya atau sumber yang bermacam-macam menjadi satu tempat.

### 3. Pemilihan Data atau Seleksi data

Merupakan pemilahan data-data yang terdapat pada database atau bisa digambarkan data yang perlu saja di ambil dari database tersebut.

### 4. Transformasi Data

Perubahan data atau disebut juga (*Data Transformation*) merupakan pengubahan pada data ataupun dicampurkan di satu database yang serasi atau sesuai untuk dilakukan proses dalam *data mining* tersebut.

### 5. Metode *Mining*

Metode mining atau biasa dikenal sebagai proses mining merupakan proses yang bertujuan atau proses yang paling utama dalam data mining untuk mencari *knowledge* (*pengetahuna baru*) dari data yang dianalisis.

### 6. Penilaian atau Evaluasi Pola Data (*Pattern Evaluation*)

Evaluasi pola atau biasa disebut (*Pattern Evaluation*) berguna mengetahui model-model atau pola yang unik ke dalam pengetahuan (*knowledge based*) yang akan dijumpai.

### 7. Pemaparan atau Presentasi Pengetahuan

Presentasi pengetahuan atau disebut (*Knowledge Presentation*) bisa disebut juga gambaran (visualisasi) serta penjabaran berhubungan dengan proses-proses yang dipakai [10].



### 2.3 Klaisfikasi

Klasifikasi adalah suatu pekerjaan yang tujuannya memperhitungkan entitas pada data tujuannya memasukan pada sebuah kelas atau variabel terpilih atas beberapa kelas yang tersedia. Didalam klasifikasi ada beberapa skema penting yang dicoba, yang pertama pengembangan modell sebagai rancangan yang di *save* atau disimpan sebagai penyimpanan serta pelaksanaan model yang bertujuan pengenalan serta prediksi atau perkiraan di objek data lain tujuannya mengetahui letak kelas atau bagian objek pada data itu sendiri kedalam model atau acuan yang sudah tersimpan [10].

Dari beberapa penjelesan dapat disimpulkan klasifikasi adalah proses untuk mendapatkan bentuk atau model fungsi tujuannya mendeskripsikan dan menjadi pembeda kelas atau bagian pada data itu sendiri, bertujuan supaya modell tersebut dapat berjalan seiring dengan memperkirakan bagian atau kelas yang tidak diketahui pada pengamatan yang dilakukan [11].

Pada masa sekarang data yang begitu besar menunjukkan bahwasannya terjadi perkembangan volume pada data sangat begitu cepat dan menghasilkan data disebut dengan (*big data*). Untuk mengantisipasinya maka diperlukan alat yang akan digunakan sebagai alat analisis yang bertujuan untuk hal-hal penting atau informasi di data yang begitu besar serta juga memperbarui atau melakukan perubahan pada data besar akan membuat pengetahuan yang terstruktur [12].

## 2.4 Naïve Bayes

Algoritma ini merupakan salah satu klasifikasi yang sering digunakan dan sederhana berguna mempertimbangkan seluruh jumlah peluang caranya memadukan beberapa gabungan serta frekuensi pada nilai basis data yang di dapatkan. Pada algoritma ini juga memanfaatkan serta memprediksi semua *variable* atau biasa disebut atribut yang leluasa dan juga saling bebas akan diperoleh nilai pada kelas atribut sendiri. *Naive Bayes* juga merupakan klasifikasi dengan prosedur atau metode peluang serta kalkulasi yang didapatkan seorang ilmuwan yang asalnya dari Negara Inggris yaitu Thomas Bayes .Dia membuat prediksi yang datang kedepannya berlandaskan hal-hal yang terdahulu [10].

*Bayes* merupakan salah satu teknik untuk memprediksi peluang sederhana berlandaskan pada implementasi teorema *bayes* dengan asumsi ketidak tergantungan yang kuat [13]. Sehingga pada implementasinya, algoritma *Naive Bayes Classifier* tidak adanya keterkaitan diantara satu variabel dengan variabel lainnya, bisa disimpulkan juga satu variabel tidak akan terjadi pengaruh dengan variabel lainnya, bahkan mungkin variabel itu sendiri saling keterkaitan [14].

*Naive Bayes* juga merupakan klasifikasi dengan prosedur atau metode peluang serta kalkulasi yang didapatkan seorang ilmuwan yang asalnya dari Negara Inggris yaitu Thomas Bayes . Beliau juga membuat prediksi yang datang kedepannya berlandaskan hal-hal yang terdahulu [15]. Berikut teorema Bayes memiliki bentuk umum sebagai berikut :

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

Di mana :

X = Kelas data yang belum diketahui

H = Hipotesa data

$P(H|X)$  = Kemungkinan Hipotesa H berdasarkan X

$P(H)$  = Kemungkinan Hipotesa H

$P(X|H)$  = Kemungkinan X berdasarkan keadaan h

$P(X)$  = Kemungkinan dari X (Persamaan 1 Teorema Naive Bayes.)

## 2.5 Probabilitas

Sedangkan klasifikasi pada probabilitas dan statistik untuk memprediksikan peluang ke depan yang akan terjadi berdasarkan dari pengalaman yang telah terjadi atau dikenal dengan *naive bayes*. untuk mencari rumus untuk mencari probabilitas yang penulis pakai adalah :

$$P = \frac{X}{N}$$

$P$  = Probabilitas /

$X$  = Jumlah Peristiwa A

$N$  = Total Peristiwa

(Persamaan 2 Rumus Probabilitas.)

## 2.6 Confusion Matrix

Matrik kebingungan atau disebut (*Confusion matrix*) adalah tabel yang disusun dengan tujuan memaparkan hasil klasifikasi dari besaran data yang akan di uji betul atau benar dari data tersebut dan salah [16]. Menurut Han dan Kamber

matrik kebingungan atau disebut juga (*Confusion matrix*) merupakan satu alat bertujuan menganalisis penggolongan itu baik dalam menandai atau mengenali urutan dengan awal mempunyai kelas yang beda. Nilai dari *True-Positive* serta nilai *True-Negative* menyajikan hal-hal penting disaat penggolongan saat melakukan klasifikasi pada data yang bernilai benar, sebaliknya *False-Positive* akan menyajikan hal-hal penting di saat penggolongan tidak benar atau salah saat melakukan klasifikasi data [17].

*Confusion matrix* juga menyajikan satu keputusan yang didapatkan pada data *training* dan *testing*, *confusion matrix* juga menyajikan nilai dari performa klasifikasi dengan landasan objek yang benar maupun tidak benar . *Confusion matrix* juga bermakna hal-hal penting yang aktual dan prediksi. Pada metode klasifikasi Dibawah ini merupakan *table* positive dan negative tuple(urutan) [18].

Tabel 2. 1 Tabel Confusion Matrix

		Predicted Class (kelas prediksi)		Total
		TRUE	FALSE	
Actual Class (kelas sebenarnya)	YES	<i>TP</i>	<i>FN</i>	<i>P</i>
	NO	<i>FP</i>	<i>TN</i>	<i>N</i>
	Total	<i>P</i>	<i>N</i>	<i>P + N</i>

**Keterangan :**

TP (True Positive) = Total data dengan nilai sebenarnya positif dan nilai perkiraan positif

FP (False Positive) = Total data dengan nilai sebenarnya negatif dan nilai perkiraan positif

FN (False Negative) = Total data dengan nilai sebenarnya positif dan nilai perkiraan negatif

TN (True Negative) = Total data dengan nilai sebenarnya negatif dan nilai perkiraan negatif

Dibawah ini merupakan beberapa model dari matrix kebingungan atau (*confusion matrix*) menurut Han & Kamber sebagai berikut :

- a) *Accuracy* (Akurasi) merupakan perbandingan atau proporsi total peluang atau prediksi nilainya benar, rumusnya :

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (\text{Persamaan 3. Rumus Akurasi}).$$

- b) *Sensitivity* atau (*Recall*) berfungsi mencari nilai peluang atau prediksi pada kasus yang kemungkinan atau di prediksi nilai positif pada faktanya tergolong pada kasus kategori positif , rumusnya:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (\text{Persamaan 4 Rumus Recall}).$$

- c) *Precision* yang berfungsi sebagai mencari angka peluang kasus atau prediksi yang kemungkinannya (kategori) bernilai positif pada keadaan faktanya tergolong pada kasus kategori yang positif , rumusnya:

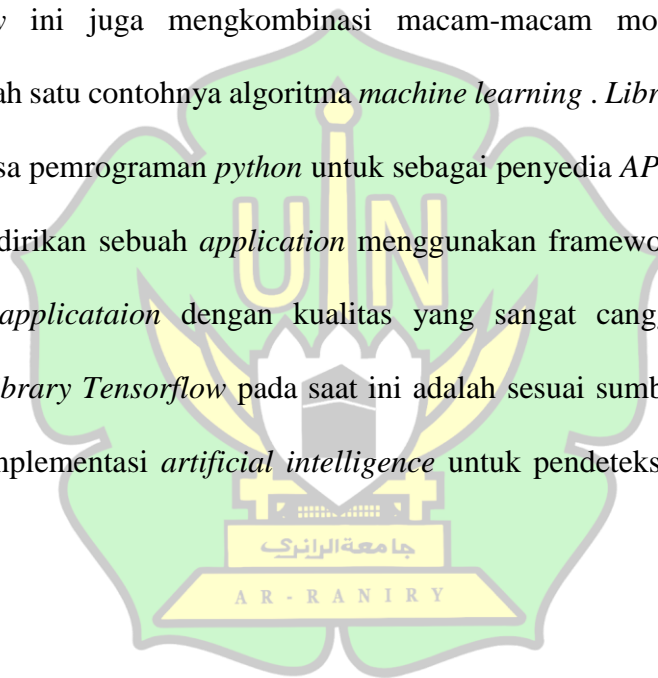
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (\text{Persamaan 5 Rumus Precision}).$$

- d) *F1-Score* kegunaannya yang berfungsi sebagai merangkum hasil dari nilai dari *sensitivity (Recall)* dan *precision* rumusnya :

$$\text{F1-Score} = \frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (\text{Persamaan 6 Rumus F1-Score}).$$

## 2.7 Tensorflow

*Tensorflow* yang diuraikan sebuah tim yaitu *Google Brain*, *Google Brain* merupakan suatu tim peneliti di perusahaan *google* yang titik fokus penelitiannya merupakan pada perkembangan (AI) kecerdasan buatan dan juga tentang *Machine Learning* pembelajaran mesin . *Tensorflow* merupakan sebuah pustaka (*library*) terbuka atau (*open library*) yang berfungsi sebagai pemrosesan angka (numerik) dengan mesin pembelajaran atau disebut (*machine learning*) dengan skala yang besar. *Library* ini juga mengkombinasi macam-macam model yang akan digunakan, salah satu contohnya algoritma *machine learning* . *Library Tensorflow* memakai bahasa pemrograman *python* untuk sebagai penyedia *API front-end* yang berfungsi mendirikan sebuah *application* menggunakan framework, bahkan juga menerapkan *applicataion* dengan kualitas yang sangat canggih. Salah satu penggunaan *Library Tensorflow* pada saat ini adalah sesuai sumber yang dipakai adalah mengimplementasi *artificial intelligence* untuk pendeteksi masker secara *real time* [19].



## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Pendekatan Penelitian**

Pada penelitian yang akan buat memakai pendekatan penelitian yaitu Kuantitatif dan jenis penelitiannya kuantitatif deskriptif. Pemilihan pendekatan ini karena penelitian ini sangat cocok karena penelitian ini yang bertujuan hanya menerapkan kinerja algoritma. Tentunya juga dengan melewati beberapa proses yang terjadi pada data mining yang berawal dari pembersihan data sampai dengan hasil yang didapatkan melalui perhitungan secara komputerisasi yang menggunakan alat bantu bahasa pemrograman *python*.

#### **3.2 Subjek Penelitian serta Sumber Data**

##### 3.2.1 Subjek Penelitian

Subjek pada penelitian yang dilakukan merujuk ke arah data penyakit Jantung yang diperoleh.

##### 3.2.2 Sumber Data

Pada penelitian ini yang menjadi yang sumber data adalah data yang di dapatkan dari sebuah *website* penyedia dataset yaitu *www.kaggle.com*.

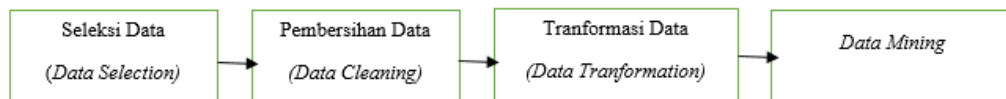
#### **3.3 Teknik Pengumpulan Data**

##### 3.3.1 Website *www.Kaggle.com*

Proses pengambilan dataset pada penelitian ini didapatkan dan diunduh dari sebuah Website yaitu *www.kaggle.com*.

### 3.4 Teknik Analisis Data

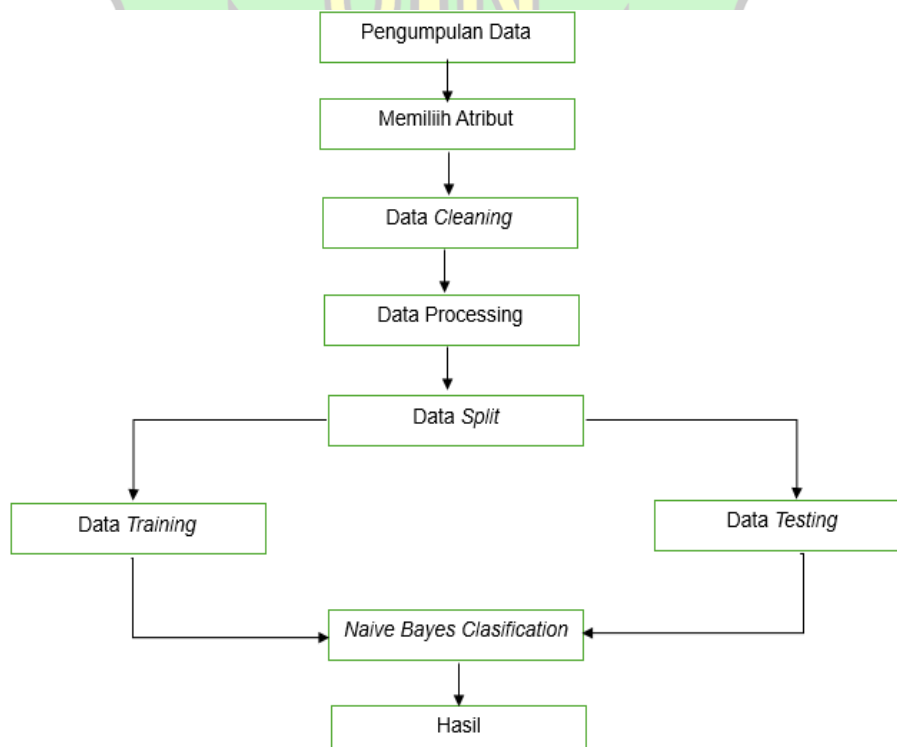
Pada penelitian yang dilakukan, untuk tahapan menganalisis data tersebut dengan proses yang bernama *data mining*. Alur-alurnya adalah sebagai gambar di bawah ini :



Gambar 3. 1 Alur Analisi Data

### 3.5 Rancangan Penelitian

Rancangan pada penelitian yang digunakan seperti flowchart di bawah ini :



Gambar 3. 2 Alur Penelitian



Penguraian dari *flowchart* penelitian :

## 1. Pengumpulan Data

Pada langkah ini akan dilaksanakannya pengumpulan-pengumpulan data , setelah didapatkan melalui sebuah website *www.kaggle.com* . Data yang dikumpulkan berkaitan dengan Data pasien yang terkena penyakit jantung yaitu ada 14 atribut yaitu , *sex* = jenis kelamin, *age* = usia, *trestbps* = tekanan darah, *cp* = angka nyeri dada, *fb* = gula darah, , *chol* = kolesterol, *resteg*= angka elektrokardiografi , *thalach* = denyut jantung, *exang* =angka angina induksi, *oldpeak* = angka despresi, *slope* = angka kemiringan segmen, *ca* = angka pembuluh darah utama, *thal* = jenis darah , *condition* – kondisi terkena . diatas adalah adalah beberapa macam atribut yang dipergunakan pada penelitian ini dan menjadi atribut dependennya adalah *condition*.

## 2. Proses Pemilihan Atribut

### a. Pemilihan Atribut

Tahapan ini berfungsi sebagai proses memilah data atau variable yang dipakai serta menghilangkan variable tidak dipergunakan. Pada kumpulan data atau (*dataset*) yang digunakan yang menjadi *data training* adalah 80% data yang telah diambil dari dataset penyakit jantung tersebut. Dan beberapa atribut yang dipakai yaitu :

### 1. Age

Dalam atribut ini umur yang rentan terkena pada 40 tahun keatas.

## 2. *Sex*

Atribut atau variabel jenis kelamin yang digunakan pada penelitian ini ada 2 dengan keterangan (perempuan = 0) dan (laki-laki = 1).

## 3. *Cp* (tipe nyeri dada)

Tipe Nyeri Dada :

- Nilai 1 : angina tipikal
- Nilai 2 : angina tipikal
- Nilai 3 : nyeri non-angina
- Nilai 4 : tanpa gejala

## 4. *Trestbps* (tekanan pada darah saat istirahat)

Tekanan pada darah disaat istirahat (dalam satuan (mm) Hgsatt terhitung masuk rumah sakit).

## 5. *Chol*

Kadar kolesterol satuan mm/Hg.

## 6. *Fbs*

Gula pada darah puasa  $> 102$  mg / dl  $\rightarrow$  ( 1 = benar, 0 = tidak).

## 7. *Restecg*

Dampak atau hasil elektrokardiografi saat istirahat.

## 8. *Thalac*

Denyut pada jantung jumlah maksimum.

## 9. *Exang*

Nyeri pada dada yang diinduksi disaat olahraga (1 = ya, 0 = salah).

10. **Oldpeak**

Depresi diakibatkan olahraga dibanding istirahat.

11. **Slope**

Kecuraman puncak latihan.

12. **Ca**

Banyak pembuluh darah utama (0-3) tandai dengan fluoroskopi.

13. **Thal**

3 = normal , 6 = cacat , tetap , 7 = cacat reversible.

14. **Condition**

Kondisi terkena penyakit jantung ( 1 = terkena , 0 = tidak terkena ).

3. **Cleaning Data atau Pembersihan Data**

Pada langkah ini dilaksanakan pembersihan pada data yang berfungsi untuk mengambil beberapa atribut yang dipakai saja untuk di terapkan pada penelitian ini dan menghilangkan *noise* pada data.

4. **Data Split**

Data Split dilakukan untuk memisahkan data menjadi 2 bagian yaitu data *training* dan data *testing* .

5. **Data Training**

Jumlah data Training yang dipakai adalah hasil pembagian dengan data testing yang dipakai berjumlah 80%, 70% dan 60%.

6. **Data Testing**

Berikut jumlah dan pembagian pada data *testing testing* yang akan dipakai:

1) Dataset x 20% Data yang diambil =  $1024 \times 20\% = 204,8 = 205$  data.

2) Dataset x 30% Data yang diambil =  $1024 \times 30\% = 307,2 = 308$  data.

3) Dataset x 40% Data yang diambil =  $1024 \times 40\% = 409.6 = 410$  data.

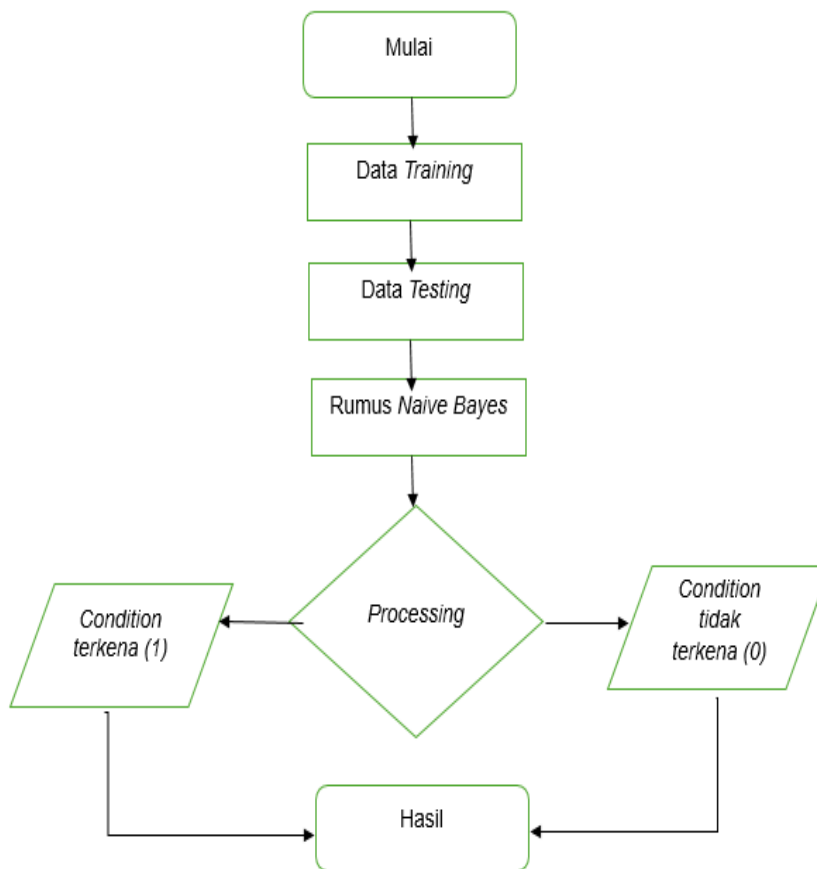
Studi empiris menunjukkan bahwa hasil terbaik diperoleh jika kita menggunakan 20-30% data untuk pengujian [20].

### 7. *Clasification Naive Bayes.*

Pada tahapan ini dari alur pertama pada penelitian akan dihitung dengan menggunakan metode *Clasification* dan akan dihitung menggunakan rumus probabilitas.

### 3.6 Alur Pengolahan Data

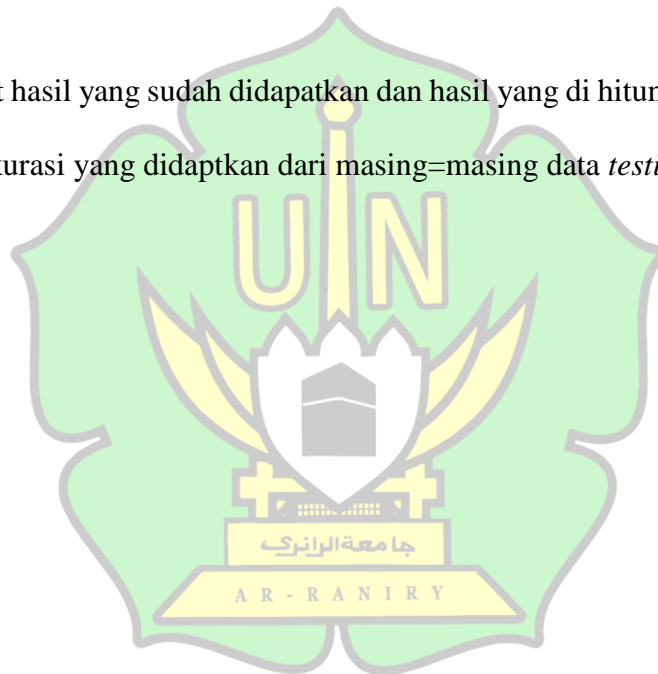
Proses atau tahap perhitungan pengolahan data :



Gambar 3. 3 Tahapan pengolahan data

Penguraian :

- a. Mempersiapkan data latih atau (*data training*) .
- b. Membagi data *testing* dari data *training* yang didapatkan
- c. Melakukan perhitungan probabilitas terhadap data *testing*
- d. Sistem membaca rumus *naive bayes* dan melakukan perhitungan dilakukan melalui sistem yang dibuat.
- e. *Processing* melalui sistem tidak secara manual dan mendapatkan hasil yang dicari.
- f. Melihat hasil yang sudah didapatkan dan hasil yang di hitung dengan sistem serta akurasi yang didapatkan dari masing-masing data *testing* yang di uji.



## BAB IV

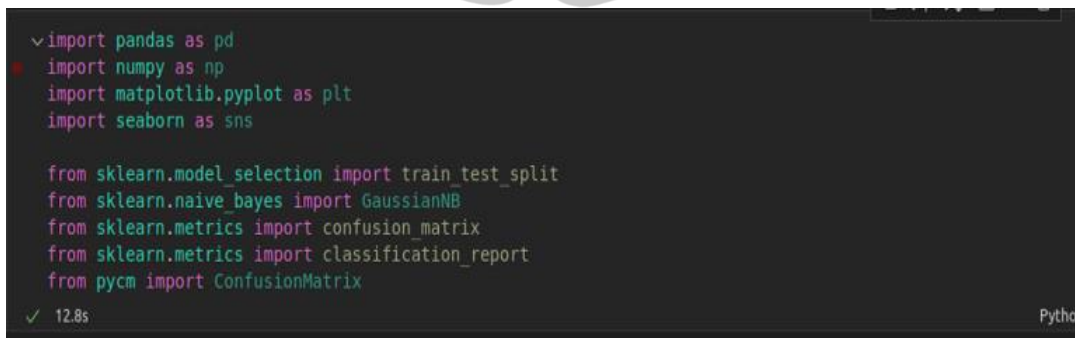
### HASIL DAN PEMBAHASAN

#### 4.1. Hasil Penelitian

##### 4.1.1. Persiapan Data

Dalam penelitian ini tahapan awal merupakan pengumpulan dan persiapan dataset yang akan digunakan untuk penelitian yang diambil dari sebuah website yang bernama *www.kaggle.com*. Data yang didapatkan berjumlah 1024 data penyakit jantung yang akan diuji dengan data testing sebesar 20% dari jumlah dataset yang didapatkan. Penelitian ini bertujuan untuk mengetahui bagaimana penerapan algoritma *Naïve Bayes* menggunakan *library Tensorflow* dan untuk mencari nilai akurasi yang didapatkan dari masing-masing data testing yang disebutkan dalam rumusan masalah sebesar 20%, 30% dan 40 % tersebut. Secara metode yang penelitian pakai dengan data *testing* yang di acak otomatis oleh sistem yang akan dipergunakan .

Dibawah ini merupakan gambar untuk memanggil beberapa didalam *library tensorflow* yang dibutuhkan tersebut untuk dipergunakan dalam penelitian ini.



```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
from pycm import ConfusionMatrix
```

✓ 12.8s Python

Gambar 4. 1 import library (sumber pribadi)

Pada tahapan pertama yang dilakukan merupakan pemanggilan beberapa *library* termasuk *library tensorflow* tersebut yang akan digunakan dalam penelitian ini sebagai langkah awal dan kemudian kita menjalankan codingan tersebut dengan perintah dibawah ini.

```
df = pd.read_csv("heart.csv")
print(df)
```

Python

Gambar 4. 2 code membaca data

Dan hasil dari codingan tersebut setelah membaca data yang diambil dari sebuah website tersebut adalah setelah dibuatkan dalam tabel sebagai berikut

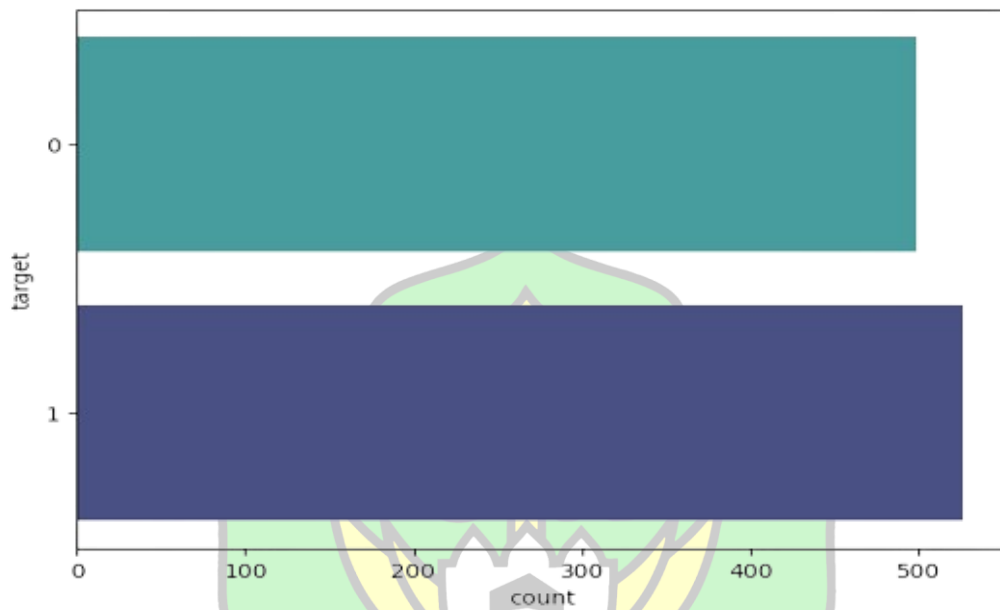
Tabel 4 1. Dataset

No	Age	Sex	Cp	Trestbps	Chol	Fbs	Resteg	Thalac	Exang	Oldpeak	Slope	Ca	Thal	target
1	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
2	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
3	78	1	0	145	174	0	1	125	1	2.6	0	0	3	0
4	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
6	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
1020	59	1	1	140	221	0	1	164	1	0.0	2	0	2	1
1021	66	1	0	125	258	0	0	141	1	2.8	1	1	3	0
1022	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
1023	50	0	0	110	245	0	0	159	0	0.0	2	0	2	1
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

Seperti gambar pada tabel diatas merupakan dataset yang sudah dijalankan diambil dari website [www.kaggle.com](http://www.kaggle.com) dengan jumlah dataset sebesar 1024 dengan atribut berjumlah 13 dan 1 atribut kunci yaitu atribut *target* yang merupakan kondisi yang dimana data tersebut dinyatakan tererekena penyakit jantung maupun tidak. Seperti yang sudah dijelaskan pada rumusan masalah diatas, penelitian ini bertujuan mencari bagaimana penerapan algoritma naïve bayes tersebut dan mencari akurasi

yang didapatkan pada algoritma dengan besar data *testing* yang digunakan adalah 20% , 30% dan 40%.

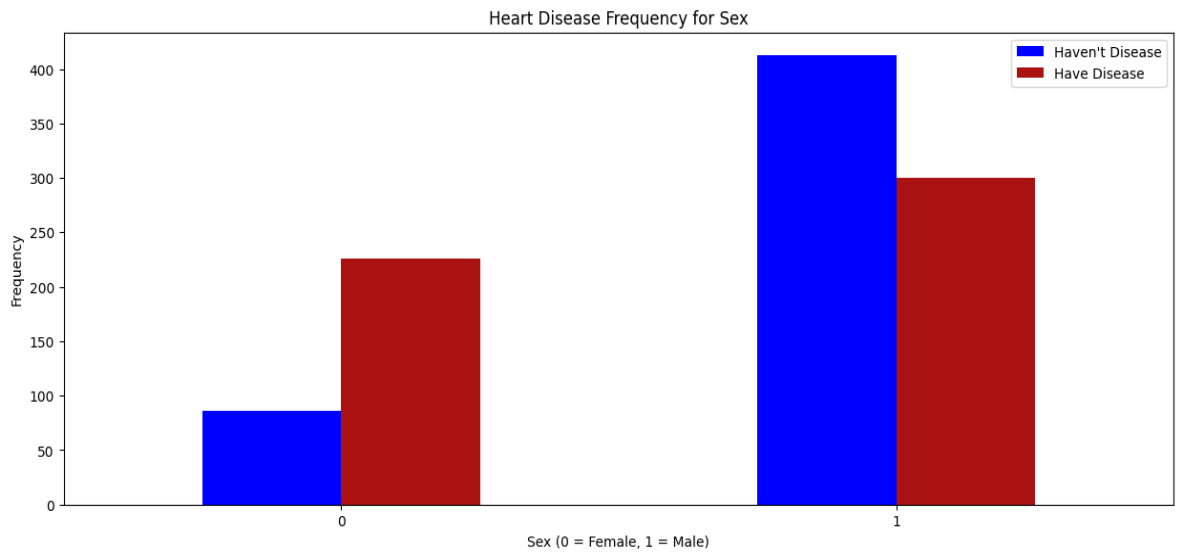
Pada data yang sudah didapatkan bisa digambarkan dalam bentuk grafik seperti dibawah ini:



Gambar 4. 3 grafik jumlah yang terkena penyakit jantung.

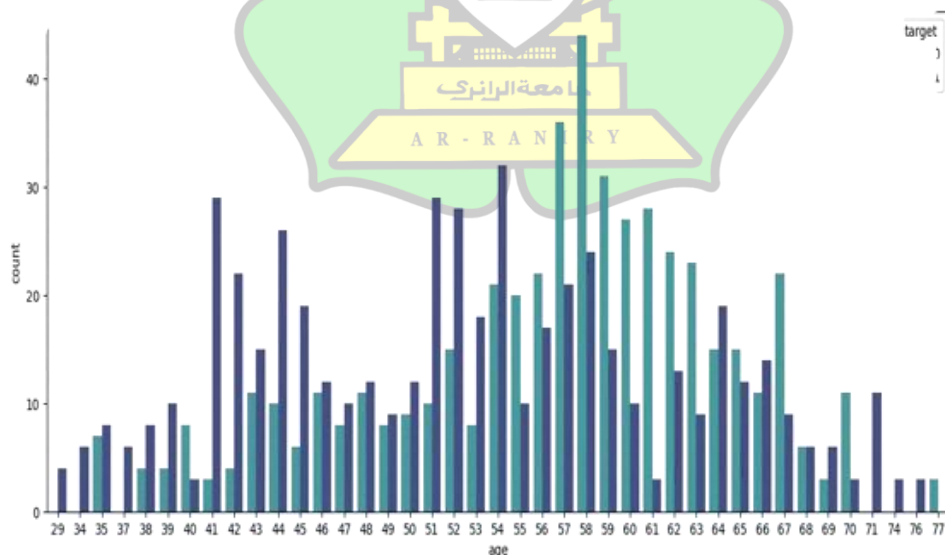
Pada gambar diatas dapat disimpulkan bahwasannya yang terkena penyakit jantung dengan (target = 1) sebanyak 526 pasien dan tidak terkena dengan (target 0) 499 pasien. Dengan frekuensi yang digambarkan juga yang diukur dari segi jenis kelamin sebagai gambar dibawah ini :





Gambar 4. 4 gambar grafik dari jenis kelamin

Penjelasan diatas merupakan frekuensi yang diukur dari jenis kelamin laki-laki dan perempuan yang terkena penyakit jantung atau tidak dengan data tertinggi terkena adalah laki-laki. Dan dibawah ini merupakan data grafik yang diukur dari segi umur seperti dibawah ini :



Gambar 4. 5 Grafik dari segi umur pasien

Penjelasan pada grafik diatas merupakan data penyakit jantung yang diukur dari segi umur pasien dari kisaran dari 29 – 77 tahun

## 4.2 Pemilihan Atribut

### 4.2.1. Pemilihan Atribut

Dalam tahapan ini akan dilakukan pemilihan atribut yang akan digunakan dari 13 dan 1 atribut kunci yang dimana atribut kunci tersebut bernama (*target*) dalam dataset yang telah diambil. Menurut pendapat yang sudah saya baca terlebih dahulu dari beberapa relevansi penelitian sebelumnya mengatakan bahwasannya algoritma *Naïve Bayes* tersebut tidak ada kaitannya antara 1 atribut dengan atribut lainnya walaupun 1 atribut tersebut saling berhubungan. Berikut atribut-atribut akan digunakan :

Tabel 4 2. Atribut

No	Column
1	<i>Age</i>
2	<i>Sex</i>
3	<i>Cp</i>
4	<i>Trestbps</i>
5	<i>Chol</i>
6	<i>Fbs</i>
7	<i>Resteg</i>
8	<i>Thalac</i>
9	<i>Exang</i>
10	<i>Olpeak</i>
11	<i>slope</i>
12	<i>Ca</i>
13	<i>Thal</i>
14	<i>Target</i>

Diatas merupakan atribut-atribut yang ada didalam dataset tersebut. pada algoritma *Naïve Bayes* atribut tidak berpengaruh ketika mencari akurasi dari algoritma tersebut karena itu penelitian ini akan menggunakan semua atribut untuk mencari akurasi dan bagaimana penerapannya dari algoritma *Naïve Bayes* tersebut.

a). Data Cleaning

Dibawah ini merupakan gambar hasil dari pembersihan data yang dilakukan pada dataset yang didapatkan

Tabel 4 3. *Cleaning data (pembersihan data)*

No	Age	Sex	Cp	Trestbps	Chol	Fbs	Resteg	Thalac	Exang	Oldpeak	Slope	Ca	Thal	target
1	false	false	false	False	false	False	false	false	false	false	false	false	false	false
2	false	false	false	False	false	False	false	false	false	false	false	false	false	false
3	false	false	false	False	false	False	false	false	false	false	false	false	false	false
4	false	false	false	False	false	False	false	false	false	false	false	false	false	false
5	false	false	false	False	false	False	false	false	false	false	false	false	false	false
6	false	false	false	False	false	False	false	false	false	false	false	false	false	false
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1020	false	false	false	False	false	False	false	false	false	false	false	false	false	false
1021	false	false	false	False	false	False	false	false	false	false	false	false	false	false
1022	false	false	false	False	false	False	false	false	false	false	false	false	false	false
1023	false	false	false	False	false	false	false	false	false	false	false	false	false	false
1024	false	false	false	False	false	false	false	false	false	false	false	false	false	false

Diatas merupakan tabel data pada dataset untuk mengetahui bahwasanya data tersebut rusak atau tidak setelah dilakukannya pembersihan data tersebut. Tujuan dari *Cleaning Data* disini adalah didalam algoritma *Naïve Bayes* ketika ada 1 data yang rusak maka akan mempengaruhi hasil yang akan diuji untuk mencari akurasi tersebut berbeda dari atribut yang tidak berpengaruh sama sekali dalam kinerja algoritma.

b). Data Split

Data split atau *splitting* merupakan langkah untuk memisahkan dua data yang bersifat data latih dan data uji . pada penelitian ini data latih yang akan di pakai sebesar 1024 data penyakit jantung dan sebagai data latih yaitu sebesar 20% , 30%

dan 40% data uji atau data *testing* dari 1024 data latih tersebut yang di ambil secara acak oleh sistem yang dibangun.

a) Data Training

Data training yang digunakan sebesar dari 1024 data penyakit jantung tersebut akan di dibagi dengan jumlah data *testing* tersebut. seperti pada gambar tabel 4.1.

b) Data testing

Data testing diambil dari dataset yang digunakan sebanyak 20% , 30% dan 40% dari banyaknya data *training* yang digunakan . dibawah ini merupakan contoh data testing dari masing-masing besaran yang diambil. Berikut adalah tabel masing-masing data *testing* yang akan digunakan dengan masing-masing jumlah yang dipakai :

Tabel 4 4. Data testing 20%

No Data	Age	Sex	Cp	Trestbps	Chol	Fbs	Resteg	Thalac	Exang	Oldpeak	Slope	Ca	Thal
85	44	1	1	120	220	0	1	170	0	0.0	2	0	<b>2</b>
210	42	1	2	120	240	1	1	194	0	0.8	0	0	<b>3</b>
182	60	1	0	140	293	0	0	170	0	1.2	1	2	<b>3</b>
872	64	1	3	170	227	0	0	155	0	0.6	1	0	<b>3</b>
716	55	0	0	128	205	0	2	138	1	2.0	1	1	<b>3</b>
....	....	....	...	.....	...	....	....	....	....	.....	....	...	...
100	43	0	2	122	213	0	1	165	0	0.2	1	0	<b>2</b>
203	64	1	3	170	227	0	0	155	0	2.6	1	0	<b>3</b>
881	57	1	1	154	232	0	0	164	0	0.0	2	1	<b>2</b>
682	59	1	0	164	176	1	0	90	0	1.0	1	2	<b>1</b>
992	50	0	0	110	254	0	0	159	0	0.0	2	0	<b>2</b>

Tabel 4 5. Variable dependen (y\_test)

No	Column
85	<i>1</i>
210	<i>1</i>
182	<i>0</i>
872	<i>1</i>
716	<i>0</i>
....	...
100	<i>1</i>
203	<i>1</i>
881	<i>0</i>
682	<i>0</i>
992	<i>1</i>

Dari tabel diatas jumlah data testing yang dipakai adalah sebesar 20% dan atribut target dari data *testing* 20%. Dibawah ini merupakan hasil dari ( $y_{test}$ ) diatas jika di gambarkan dalam bentuk array sebelum dilakukannya prediksi dari algoritma *naive bayes* dan sesudah dilakukan prediksi dengan keterangan sebagai berikut :

- a.  $X_{train}$  merupakan data yang akan kita latih yaitu sebanyak 1024 dataset dibagikan dengan data *testing* yang ditetapkan.
- b.  $X_{test}$  merupakan data yang akan kita test yaitu data *testing* sebanyak 20% , 30% dan 40%.
- c.  $Y_{train}$  merupakan variable dependen yang dilatih dan telah dipisahkan yaitu variable target yang di dataset (terkena atau tidak terkena penyakit jantung).
- d.  $Y_{test}$  merupakan vaiabel dependen yang diuji yaitu tempat variable target tersebut yang akan diuji dengan  $Y_{train}$  tersebut.

```

np.array(y_test)
[24] array([[1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1,
1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0,
0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1,
0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1,
0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0,
0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1,
0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
1, 1, 1, 1, 0, 0, 1])

```

Gambar 4. 6 (y\_test) atribut target sebelum di prediksi

```

# Menentukan hasil prediksi dari x_test
y_pred = nbtrain.predict(x_test)
y_pred
[29] array([[1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1,
1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1,
0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1,
0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0,
0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0,
1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
1, 1, 1, 1, 1, 0, 1])

```

Gambar 4. 7 (y\_test) aatribut target yang sudah diprediksi

Pada hasil yang sudah diprediksi diatas akan dimasukkan kedalam *confusion matrix* untuk mencari akurasi kinerja dari algoritma tersebut. Dibawah ini merupakan tabel dari data *testing* sebesar 30% :

Tabel 4.6. Data Testing 30%

No Data	Age	Sex	Cp	Trestbps	Chol	Fbs	Resteg	Thalac	Exang	Oldpeak	Slope	Ca	Thal
85	44	1	1	120	220	0	1	170	0	0.0	2	0	2
210	42	1	2	120	240	1	1	194	0	0.8	0	0	3
182	60	1	0	140	293	0	0	170	0	1.2	1	2	3
872	64	1	3	170	227	0	0	155	0	0.6	1	0	3
716	55	0	0	128	205	0	2	138	1	2.0	1	1	3
....	....	....	....	.....	....	....	....	....	....	.....	....	....	....
343	52	1	2	172	199	1	1	162	0	0.5	2	0	3
669	61	1	0	138	166	0	0	125	1	3.6	1	1	2
237	57	1	1	124	261	0	1	141	0	0.3	2	0	3
461	48	1	2	124	255	1	1	175	0	0.0	2	2	2
316	64	0	2	140	313	0	1	133	0	0.2	2	0	3

Tabel 4 7. Variable dependen (*y\_test*)

No	Column
85	<i>1</i>
210	<i>1</i>
182	<i>0</i>
872	<i>1</i>
716	<i>0</i>
....	...
343	<i>1</i>
660	<i>0</i>
237	<i>0</i>
461	<i>1</i>
316	<i>1</i>

Dari tabel diatas jumlah data testing yang dipakai adalah sebesar 30% dan atribut target dari data *testing* 30%. Dibawah ini merupakan hasil dari (*y\_test*) diatas jika di gambarkan dalam bentuk array sebelum dilakukannya prediksi dari algoritma *naive bayes* dan sesudah dilakukan prediksi:

```

np.array(y_test)
[18] Python
... array([1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1,
1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0,
0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1,
0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1,
0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0,
0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1,
0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1,
0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1,
0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0,
0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0,
1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1])

```

Gambar 4. 8 (*y\_test*) atribut target sebelum di prediksi

```

# Menentukan hasil prediksi dari x_test
y_pred = nbtrain.predict(x_test)
y_pred

[17] Python
... array([1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1,
1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1,
0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1,
0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0,
0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0,
1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1,
1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0,
0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0,
0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1,
0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0,
0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 1])

```

Gambar 4. 9 (y\_test) atribut target sesudah di prediksi

Pada hasil yang sudah diprediksi diatas akan dimasukkan kedalam *confusion matrix* untuk mencari akurasi kinerja dari algoritma tersebut. Dan berikutnya adalah tabel data *testing* yang berjumlah 40% dari 1024 dataset :

Tabel 4 8. Data Testing 40%

No Data	Age	Sex	Cp	Trestbps	Chol	Fbs	Resteg	Thalac	Exang	Oldpeak	Slope	Ca	Thal
85	44	1	1	120	220	0	1	170	0	0.0	2	0	2
210	42	1	2	120	240	1	1	194	0	0.8	0	0	3
182	60	1	0	140	293	0	0	170	0	1.2	1	2	3
872	64	1	3	170	227	0	0	155	0	0.6	1	0	3
716	55	0	0	128	205	0	2	138	1	2.0	1	1	3
....	....	....	....	....	....	....	....	....	....	....	....	....	....
746	64	1	2	140	335	0	1	158	0	0.0	2	0	2
595	61	1	0	148	203	0	1	161	0	0.0	2	1	3
160	77	1	0	125	304	0	0	162	1	0.0	2	3	2
976	63	0	0	108	269	0	1	169	1	1.0	1	2	2
120	54	1	1	108	309	0	1	156	0	0.0	2	0	3



Tabel 4 9. Variable dependen (y\_test)

No	Column
85	1
210	1
182	0
872	1
716	0
....	...
746	0
595	0
160	0
976	0
120	1

Dari tabel diatas jumlah data testing yang dipakai adalah sebesar 30% dan atribut target dari data *testing* 30%. Dibawah ini merupakan hasil dari (y\_test) diatas jika di gambarkan dalam bentuk array sebelum dilakukannya prediksi dari algoritma *naive bayes* dan sesudah dilakukan prediksi:

```

np.array(y_test)
[23] Python
... array([1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1,
1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0,
0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1,
0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1,
0, 1, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0,
0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0,
1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1,
0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1,
0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0,
0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0,
1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1,
1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1,
1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1,
0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0,
0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1,
0, 1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1])

```

Gambar 4. 10 (y\_test) atribut target sebelum di prediksi



### 3) Rumus *Naive Bayes*

Pada tahapan ini sistem , rumus tersebut dirancang dan disesuaikan oleh sistem, berikut rumus yang dipakai

```
#DATA TEST DIAMBIL 20%

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state = 123)

modelnb = GaussianNB()

nbtrain = modelnb.fit(x_train, y_train)
```

Gambar 4. 12. rumus *naive bayes*

#### d) Hasil

Berikut adalah tabel hasil akurasi dari confusion matrix yang telah di uji :

Tabel 4 10. Cofusion matrix data testing 20%

		Prediction	
		1	0
Actual	0	85	19
	1	14	87

Accuracy

$$= \frac{TP + TN}{TP + TN + FP + FN}$$

$$= \frac{85 + 87}{85 + 87 + 19 + 14}$$

$$\frac{172}{205} = 0.83$$

Sensitivity (recall) =  $\frac{TP}{TP + FN}$

$$= \frac{85}{85 + 14}$$

$$\frac{85}{99} = 0,85$$

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP + FP} \\ &= \frac{85}{85 + 19} \\ &= \frac{85}{104} = 0.81 \end{aligned}$$

$$\begin{aligned} \text{F1-Score} &= \frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})} \\ &= \frac{2 \times (0.85 \times 0.81)}{(0.85 + 0.81)} \\ &= \frac{1,377}{1,66} = 0.82 \end{aligned}$$

dengan hasil akurasi yang didapatkan adalah sebagai berikut :

Tabel 4 11. akurasi dari 20%

<i>Target</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
	0.81	0,85	0,82	
<i>Accuracy</i>	0.83			<b>205</b>
<i>Macro avg</i>	0.84	0.84	0.84	<b>205</b>
<i>Weighted avg</i>	0.84	0.84	0.84	<b>205</b>

Tabel 4 12. confusion matrix data testing 30%

<b>Prediction</b>	<b>1</b>	<b>0</b>
<b>Actual</b>		
<b>0</b>	128	<b>26</b>
<b>1</b>	23	<b>131</b>

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\frac{128 + 131}{128 + 131 + 26 + 23}$$

$$\frac{259}{308} = 0.84$$

Sensitivity (recall) =  $\frac{TP}{TP + FN}$

$$\frac{128}{128 + 23}$$

$$\frac{128}{151} = 0,84$$

Precesion =  $\frac{TP}{TP + FP}$

$$\frac{128}{128 + 26}$$

$$\frac{128}{154} = 0,83$$

F1-Score =  $\frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})}$

$$\frac{2 \times (0.84 \times 0.83)}{(0.84 + 0.83)}$$

$$\frac{2,944}{1,67} = 0.83$$

$$1,3944 = 0.83$$

$$1,67$$

dengan hasil akurasi yang didapatkan adalah sebagai berikut :

Tabel 4 13. akurasi dari 30%

<i>Target</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
	0.83	0.84	0.83	
<i>Accuracy</i>	0.84			<b>308</b>
<i>Macro avg</i>	0.84	0.84	0.84	<b>308</b>
<i>Weighted avg</i>	0.84	0.84	0.84	<b>308</b>

Tabel 4 14. confusion matrix data testing 40%

Prediction Actual	1	0
0	166	42
1	33	169

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\frac{166 + 169}{166 + 169 + 42 + 33}$$

$$\frac{335}{410} = 0.81$$

$$\text{Sensitivity (recall)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\frac{166}{166 + 33}$$

$$\frac{166}{199} = 0,83$$

$$\text{Precesion} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\frac{166}{166 + 42}$$

$$\frac{166}{208} = 0.79$$

$$\text{F1-Score} = \frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})}$$

$$\frac{2 \times (0.83 \times 0.79)}{(0.83 + 0.79)}$$

$$\frac{1,3114}{1,62} = 0.80$$

dengan hasil akurasi yang didapatkan adalah sebagai berikut :

Tabel 4 15. akurasi data testing 40%

<i>Target</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
	0,79	0,83	0,80	
<i>Accruracy</i>	0,81			<b>410</b>
<i>Macro avg</i>	0.82	0.82	0.82	<b>410</b>
<i>Weighted avg</i>	0.82	0.82	0.82	<b>410</b>

Dari tabel-tabel diatas maka dapat dilihat seperti dibawah ini :

Tabel 4 16. hasil akhir

<b>Data Testing</b>	<b>20%</b>	<b>30%</b>	<b>40%</b>
<i>Accuracy</i>	0,83	0.84	0.82
<i>Precision</i>	0,81	0,83	0,79
<i>Recall</i>	0,85	0,84	0,83
<i>F1-Score</i>	0,82	0,83	0,80

Hasil akhir dari penelitian ini adalah masing-masing dari setiap data *testing* yang diuji adalah mendapatkan dari *accuracy* , *precision* , *recall*, *f1-score* dengan berbeda-beda. Dan dapat disimpulkan akurasi yang paling tinggi terdapat pada data *testing* 30% sebesar 0,84%.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

- 1) Mengetahui bagaimana perhitungan yang terjadi pada algoritma *Naive Bayes* tersebut dengan menghitung antara variable-variable lainnya dan memprediksi terkena penyakit jantung apa tidak dari data yang terdahulu.
- 2) Setelah dilakukan penelitian ini hasil yang didapatkan dengan menggunakan algoritma *Naive Bayes* dengan data data *testing* yang sudah di tetapkan secara berturut-turut sebesar 83%, 84% dan 82% dan hasil tersebut lebih besar dari penelitian yang dilakukan sebelum nya menggunakan algoritma *Decision Tree* dengan akurasi sebesar 80%

#### 5.2 Saran

Pada penelitian ini penulis menyarankan untuk kedepannya mencari *accuracy* dengan metode lainnya dengan data testing yang sama untuk mengetahui tingkat *accuracy* dari algoritma tersebut.



## DAFTAR PUSTAKA

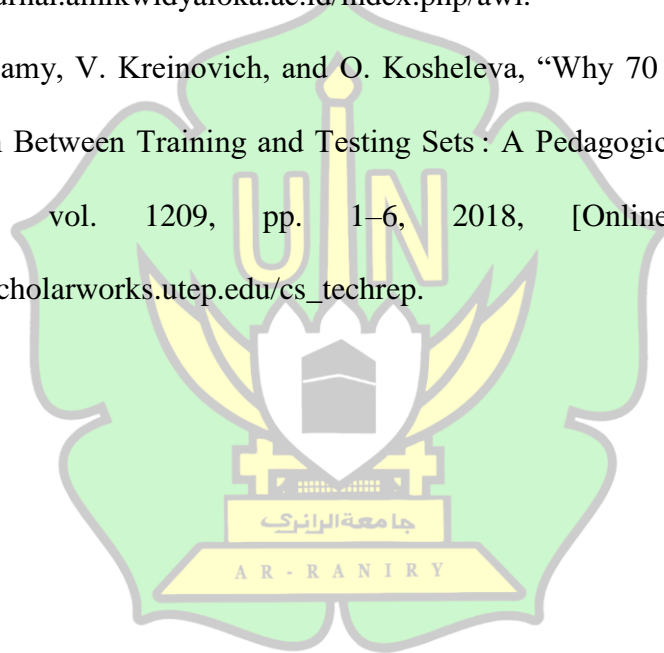
- [1] I. M. Agus Oka Gunawan, I. D. A. Indah Saraswati, I. D. G. Riswana Agung, and I. P. Eka Putra, “Klasifikasi Penyakit Jantung Menggunakan Algoritma Decision Tree Series C4.5 Dengan Rapidminer,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 2, pp. 73–83, 2023, doi: 10.47233/jteksis.v5i2.775.
- [2] Syarli and A. A. Muin, “Metode Naive Bayes Untuk Prediksi Kelulusan,” *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2020, [Online]. Available: <https://media.neliti.com/media/publications/283828-metode-naive-bayes-untuk-prediksi-kelulu-139fcfea.pdf>.
- [3] M. Melyani, L. N. Tambunan, and E. P. Baringbing, “Hubungan Usia dengan Kejadian Penyakit Jantung Koroner pada Pasien Rawat Jalan di RSUD dr. Doris Sylvanus Provinsi Kalimantan Tengah,” *J. Surya Med.*, vol. 9, no. 1, pp. 119–125, 2023, doi: 10.33084/jsm.v9i1.5158.
- [4] D. Larassati, A. Zaidiah, and S. Afrizal, “Sistem Prediksi Penyakit Jantung Koroner Menggunakan Metode Naive Bayes,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 7, no. 2, pp. 533–546, 2022, doi: 10.29100/jipi.v7i2.2842.
- [5] E. W. Fridayanthie, “ANALISA DATA MINING UNTUK PREDIKSI PENYAKIT HEPATITIS DENGAN MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE,” *J. KHATULISTIWA Inform.*, vol. 3, no. 1, 2015.
- [6] J. R. S. Penda Sudarto Hasugian, “Penerapan Data Mining Untuk

- Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Algoritma K-Means,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 3, pp. 262–268, 2022, [Online]. Available: <https://djournals.com/klik>.
- [7] A. Mardiaha and Y. Yulia, “Implementasi Data Mining Menggunakan Algoritma Apriori Pada Penjualan Suku Cadang Motor,” *J. Ilmu Komput.*, vol. 14, no. 2, p. 125, 2021, doi: 10.24843/jik.2021.v14.i02.p07.
- [8] N. Manullang, R. W. Sembiring, I. Gunawan, I. Parlina, and I. Irawan, “Implementasi Teknik Data Mining untuk Prediksi Peminatan Jurusan Siswa Menggunakan Algoritma C4.5,” *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 1–5, 2021, doi: 10.35960/ikomti.v2i2.700.
- [9] E. T. Naldy and A. Andri, “Penerapan Data Mining Untuk Analisis Daftar Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Transaksi Penjualan Toko Bangunan MDN,” *J. Nas. Ilmu Komput.*, vol. 2, no. 2, pp. 89–101, 2021, doi: 10.47747/jurnalnik.v2i2.525.
- [10] M. Sabransyah, Y. N. Nasution, and D. Tisna, “Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung,” *J. EKSPONENSIAL*, vol. 8, no. 2, pp. 111–118, 2017, [Online]. Available: <http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/31>.
- [11] P. R. Sihombing and I. F. Yuliati, “Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia,” *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 2, pp. 417–426, 2021, doi: 10.30812/matrik.v20i2.1174.
- [12] A. Merdekawati and J. T. Kumalasari, “Komparasi Dua Metode Algoritma


- Klasifikasi Untuk Prediksi Pemberian Kartu Jakarta Pintar,” *InfoTekJar J. Nas.* ..., vol. 2, 2022, [Online]. Available: [https://repository.bsi.ac.id/repo/files/342320/download/Jurnal-ATD\\_JTK.pdf](https://repository.bsi.ac.id/repo/files/342320/download/Jurnal-ATD_JTK.pdf).
- [13] L. M. Budi Harijanto, Yuri Ariyanto, “PENETAPAN ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI RETENSI ARSIP,” *Junal Inform. Polinema*.
- [14] A. Z. M. Sigid Widodo, A. Pandu Kusuma, and W. Dwi Puspitasari, “Analisis Algoritma Naive Bayes Classifier (Nbc) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 7, no. 1, pp. 87–94, 2023, doi: 10.36040/jati.v7i1.5692.
- [15] 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.
- [16] D. Normawati and S. A. Prayogi, “Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *J. Sains Komput. Inform. (J-SAKTI,* vol. 5, no. 2, pp. 697–711, 2021, [Online]. Available: <https://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369/348>.
- [17] M. Fluorida Fibrianda and A. Bhawiyuga, “Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naive Bayes Dan Support Vector Machine (SVM),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.,* vol. 2, no. 9, pp. 3112–3123, 2018, [Online]. Available: <http://j->

ptiik.ub.ac.id.

- [18] 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.
- [19] R. F. Muharram and A. Suryadi, "Implementasi artificial intelligence untuk deteksi masker secara realtime dengan tensorflow dan ssdmobilenet Berbasis python," *J. Widya*, vol. 3, no. 2, pp. 281–290, 2022, [Online]. Available: <https://jurnal.amikwidyaloka.ac.id/index.php/awl>.
- [20] A. Gholamy, V. Kreinovich, and O. Kosheleva, "Why 70 / 30 or 80 / 20 Relation Between Training and Testing Sets : A Pedagogical," *Dep. Tech. Reports*, vol. 1209, pp. 1–6, 2018, [Online]. Available: [https://scholarworks.utep.edu/cs\\_techrep](https://scholarworks.utep.edu/cs_techrep).



## Lampiran 1 Surat Keterangan Skripsi

  
**KEPUTUSAN DEKAN FAKULTAS TARBİYAH DAN KEGURUAN UIN AR-RANIRY BANDA ACEH**  
NOMOR: B-2342/Un.08.FTK/Kp.07.6/3/2024

**TENTANG:**  
**PENGANGKATAN PEMBIMBING SKRIPSI MAHASISWA**  
**DENGAN RAHMAT TUHAN YANG MAHA ESA**

**DEKAN FAKULTAS TARBİYAH DAN KEGURUAN UIN AR-RANIRY BANDA ACEH**

**Menimbang** :

- bahwa untuk kelancaran bimbingan skripsi mahasiswa pada Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh maka dipandang perlu menunjuk pembimbing skripsi,
- bahwa yang namanya tersebut dalam Surat Keputusan ini dianggap cakap dan mampu untuk diangkat dalam jabatan sebagai pembimbing skripsi mahasiswa,
- bahwa berdasarkan pertimbangan sebagaimana dimaksud dalam huruf a dan huruf b, perlu menetapkan Keputusan Dekan Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh.

**Mengingat** :

- Undang-Undang Nomor 20 Tahun 2003, tentang Sistem Pendidikan Nasional;
- Undang-Undang Nomor 14 Tahun 2005, tentang Guru dan Dosen;
- Undang-Undang Nomor 12 Tahun 2012, tentang Pendidikan Tinggi;
- Peraturan Presiden Nomor 74 Tahun 2012, tentang perubahan atas peraturan pemerintah RI Nomor 23 Tahun 2005 tentang pengelolaan keuangan Badan Layanan Umum;
- Peraturan Pemerintah Nomor 4 Tahun 2014, tentang penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi;
- Peraturan Presiden Nomor 64 Tahun 2013, tentang perubahan Institut Agama Islam Negeri Ar-Raniry Banda Aceh Menjadi Universitas Islam Negeri Ar-Raniry Banda Aceh;
- Peraturan Menteri Agama RI Nomor 44 Tahun 2022, tentang Organisasi dan Tata Kerja UIN Ar-Raniry Banda Aceh;
- Peraturan Menteri Agama Nomor 14 Tahun 2022 tentang Statuta UIN Ar-Raniry Banda Aceh;
- Keputusan Menteri Agama Nomor 492 Tahun 2003, tentang Pendelegasian Wewenang Pengangkatan, Pemindahan dan Pemberhentian PNS di Lingkungan Depag RI;
- Keputusan Menteri Keuangan Nomor 293/Kmk/05/2011, tentang penetapan UIN Ar-Raniry Banda Aceh pada Kementerian Agama sebagai Instansi Pemerintah yang menerapkan Pengelolaan Badan Layanan Umum;
- Surat Keputusan Rektor UIN Ar-Raniry Banda Aceh Nomor 01 Tahun 2015, tentang Pendelegasian Wewenang kepada Dekan dan Direktur Pascasarjana di Lingkungan UIN Ar-Raniry Banda Aceh.

**MEMUTUSKAN**

**Menetapkan** : Keputusan Dekan Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh tentang Pembimbing Skripsi Mahasiswa.

**KESATU** : Menunjukkan Saudara :  
**Nazaruddin Ahmad, M.T.**  
Untuk membimbing Skripsi

Nama : Muhammad Risyad Rahmadi  
NIM : 200212062  
Program Studi : Pendidikan Teknologi Informasi  
Judul Skripsi : Penerapan Algoritma Naive Bayes Untuk Prediksi Penyakit Jantung Menggunakan Library Tensorflow


**EDUA** : Kepada pembimbing yang tercantum namanya diatas diberikan honorarium sesuai dengan peraturan perundang-undangan yang berlaku;

**KETIGA** : Pembiayaan akibat keputusan ini dibebankan pada DIPA UIN Ar-Raniry Banda Aceh Nomor SP DIPA-025.04.2.423925/2024 Tanggal 24 November 2023 Tahun Anggaran 2024;

**KEEMPAT** : Surat Keputusan ini berlaku selama enam bulan sejak tanggal ditetapkan;


**KELIMA** : Surat Keputusan ini berlaku sejak tanggal ditetapkan dengan ketentuan bahwa segala sesuatu akan diubah dan diperbaiki kembali sebagaimana mestinya, apabila kemudian hari ternyata terdapat kekeliruan dalam Surat Keputusan ini.

Ditetapkan di : Banda Aceh  
pada tanggal : 04 Maret 2024

  
Safri Sulukh

**Tembusan**

- Sekjen Kementerian Agama RI di Jakarta;
- Dirjen Pendidikan Islam Kementerian Agama RI di Jakarta;
- Direktur Perguruan Tinggi Agama Islam Kementerian Agama RI di Jakarta;
- Kantor Pelayanan Perbendaharaan Negara (KPPN), di Banda Aceh;
- Rektor UIN Ar-Raniry Banda Aceh di Banda Aceh;
- Yachtia Bagian Keuangan dan Akuntansi UIN Ar-Raniry Banda Aceh di Banda Aceh;
- Yang Berkepentingan;
- Asas



## Lampiran 2 Form Bimbingan

### FORM KONSULTASI PEMBIMBINGAN PENULISAN SKRIPSI

Nama : Muhammad Risyad Rahmadi  
 NIM : 200212062  
 Program Studi : Pendidikan Teknologi Informasi  
 Judul Skripsi/Tugas akhir : PENERAPAN ALGORITMA NAÏVE BAYES UNTUK  
 PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN  
 LIBRARY TENSORFLOW  
 Pembimbing : Nazaruddin Ahmad, M.T.

No	Tanggal	Topik/Bab	Saran Pembimbing	Paraf
1.	14-03-2024	Pembuatan bab I.	Pembuatan awal bab I.	M. Nur
2.	18-03-2024	Pembuatan bab I	Revisi bab II.	M. Nur
3.	26-03-2024	Pembuatan bab I	Acc bab I	M. Nur
4.	13-05-2024	Pembuatan bab II	Pembuatan awal bab II.	M. Nur
5.	17-05-2024	Pembuatan bab II	Revisi bab II	M. Nur
6.	20-05-2024	Pembuatan bab II	Revisi bab II	M. Nur
7.	28-05-2024	Akhir bab II	Acc bab II	M. Nur
8.	6-06-2024	Pembuatan bab III	Pembuatan awal bab III.	M. Nur
9.	10-06-2024	Pembuatan Bab III	Revisi Bab III.	M. Nur
10.	24-06-2024	Pembuatan Bab IV.	Revisi bab IV.	M. Nur
11.	28-06-2024	Pembuatan Bab IV	Acc Bab IV.	M. Nur
12.	30-06-2024	Pembuatan Bab V	Acc Bab V.	M. Nur

Mengetahui,

Dosen Pembimbing

*M. Nur*

Nazaruddin Ahmad, M.T.  
 NIP/NIDN.0105068202

### Lampiran 3 Data Testing

no data	age	sex	cp	terstbps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
85	44	1	1	120	220	0	1	170	0	0.0	2	0	2
210	42	1	2	120	240	1	1	194	0	0.8	0	0	3
182	60	1	0	140	293	0	0	170	0	1.2	1	2	3
872	64	1	3	170	227	0	0	155	0	0.6	1	0	3
716	55	0	0	128	205	0	2	130	1	2.0	1	1	3
887	57	1	0	110	201	0	1	126	1	1.5	1	0	1
657	39	0	2	138	220	0	1	152	0	0.0	1	0	2
1017	53	1	0	123	282	0	1	95	1	2.0	1	2	3
821	62	0	0	140	268	0	0	160	0	3.6	0	2	2
813	53	1	0	140	203	1	0	155	1	3.1	0	0	3
375	66	1	0	160	228	0	0	138	0	2.3	2	0	1
280	45	0	1	130	234	0	0	175	0	0.6	1	0	2
816	70	1	1	156	245	0	0	143	0	0.0	2	0	2
970	38	1	2	138	175	0	1	173	0	0.0	2	4	2
920	39	1	0	118	219	0	1	140	0	1.2	1	0	3
511	43	1	0	120	177	0	0	120	1	2.5	1	0	3
338	56	1	2	130	256	1	0	142	1	0.6	1	1	1
221	63	1	0	130	254	0	0	147	0	1.4	1	1	3
353	57	1	0	110	201	0	1	126	1	1.5	1	0	1
890	55	0	0	128	205	0	2	130	1	2.0	1	1	3
156	40	1	3	140	199	0	1	178	1	1.4	2	0	3
167	57	0	0	120	354	0	1	163	1	0.6	2	0	2
935	56	1	1	120	240	0	1	169	0	0.0	0	0	2
134	54	1	1	108	309	0	1	156	0	0.0	2	0	3
209	62	1	1	120	281	0	0	103	0	1.4	1	1	3
747	60	1	0	117	230	1	1	160	1	1.4	2	2	3
196	58	1	0	100	234	0	1	156	0	0.1	2	1	3
835	49	1	2	118	149	0	0	126	0	0.8	2	3	2
711	35	1	0	120	198	0	1	130	1	1.6	1	0	3
172	56	1	1	120	240	0	1	169	0	0.0	0	0	2
13	51	1	0	140	298	0	1	122	1	4.2	1	3	3
272	39	0	2	138	220	0	1	152	0	0.0	1	0	2
98	65	1	0	110	248	0	0	158	0	0.6	2	2	1
909	50	1	0	144	200	0	0	126	1	0.9	1	0	3
900	61	1	3	134	234	0	1	145	0	2.6	1	2	2
838	51	1	3	125	213	0	0	125	1	1.4	2	1	2
285	71	0	2	110	265	1	0	130	0	0.0	2	1	2
800	67	1	0	120	229	0	0	129	1	2.6	1	2	3
426	54	0	2	160	201	0	1	163	0	0.0	2	1	2
348	43	1	0	132	247	1	0	143	1	0.1	1	4	3
171	56	1	0	130	283	1	0	103	1	1.6	0	0	3
274	66	1	0	160	228	0	0	138	0	2.3	2	0	1

no data	age	sex	cp	terstbpps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
859	51	1	0	140	299	0	1	173	1	1.6	2	0	3
456	47	1	2	108	243	0	1	152	0	0.0	2	0	2
845	56	1	0	132	184	0	0	105	1	2.1	1	1	1
7	55	1	0	160	289	0	0	145	1	0.8	1	1	3
95	45	0	0	138	236	0	0	152	1	0.2	1	0	2
396	68	1	2	180	274	1	0	150	1	1.6	1	0	3
336	57	1	2	150	126	1	1	173	0	0.2	2	1	3
372	54	0	2	110	214	0	1	158	0	1.6	1	0	2
344	41	1	1	120	157	0	1	182	0	0.0	2	0	2
1020	59	1	1	140	221	0	1	164	1	0.0	2	0	2
534	54	0	2	108	267	0	0	167	0	0.0	2	0	2
249	42	1	2	130	180	0	1	150	0	0.0	2	0	2
425	51	0	0	130	305	0	1	142	1	1.2	1	0	3
662	47	1	0	112	204	0	1	143	0	0.1	2	0	2
50	58	0	3	150	283	1	0	162	0	1.0	2	0	2
389	63	1	3	145	233	1	0	150	0	2.3	0	0	1
975	39	1	0	118	219	0	1	140	0	1.2	1	0	3
43	46	1	0	120	249	0	0	144	0	0.8	2	0	3
648	71	0	0	112	149	0	1	125	0	1.6	1	0	2
792	68	1	0	144	193	1	1	141	0	3.4	1	2	3
114	59	1	2	126	218	1	1	134	0	2.2	1	1	1
457	39	0	2	138	220	0	1	152	0	0.0	1	0	2
188	63	1	0	130	254	0	0	147	0	1.4	1	1	3
470	60	0	3	150	240	0	1	171	0	0.9	2	0	2
655	41	1	1	110	235	0	1	153	0	0.0	2	0	2
318	59	1	0	140	177	0	1	162	1	0.0	2	1	3
261	54	1	2	150	232	0	0	165	0	1.6	2	0	3
603	63	0	1	140	195	0	0	179	0	0.0	2	2	2
684	60	1	2	140	185	0	0	155	0	3.0	1	0	2
298	35	1	1	122	192	0	1	174	0	0.0	2	0	2
270	43	1	0	110	211	0	1	161	0	0.0	2	0	3
666	35	1	1	122	192	0	1	174	0	0.0	2	0	2
1019	47	1	0	112	204	0	1	143	0	0.1	2	0	2
147	41	1	0	110	172	0	0	158	0	0.0	2	0	3
273	54	1	0	110	206	0	0	108	1	0.0	1	1	2
654	41	0	1	126	306	0	1	163	0	0.0	2	0	2
785	55	1	0	160	289	0	0	145	1	0.8	1	1	3
643	65	1	0	120	177	0	1	140	0	0.4	2	0	3
934	42	1	2	130	180	0	1	150	0	0.0	2	0	2
915	46	1	0	120	249	0	0	144	0	0.8	2	0	3
956	49	0	0	130	269	0	1	163	0	0.0	2	0	2
388	64	1	0	120	246	0	0	96	1	2.2	0	1	2
556	59	1	0	140	177	0	1	162	1	0.0	2	1	3



no data	age	sex	cp	terstbpps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
523	61	1	0	148	203	0	1	161	0	0.0	2	1	3
335	58	1	0	150	270	0	0	111	1	0.8	2	0	3
536	50	0	2	120	219	0	1	158	0	1.6	1	0	2
999	67	1	0	125	254	1	1	163	0	0.2	1	2	3
448	51	0	2	120	295	0	0	157	0	0.6	2	0	2
694	39	1	0	118	219	0	1	140	0	1.2	1	0	3
852	52	1	0	128	255	0	1	161	1	0.0	2	1	3
710	56	1	1	130	221	0	0	163	0	0.0	2	0	3
303	60	1	0	145	282	0	0	142	1	2.8	1	2	3
798	59	1	3	170	288	0	0	159	0	0.2	1	0	3
374	46	0	2	142	177	0	0	160	1	1.4	0	0	2
164	56	1	0	125	249	1	0	144	1	1.2	1	1	2
1006	58	1	2	140	211	1	0	165	0	0.0	2	0	2
103	47	1	2	138	257	0	0	156	0	0.0	2	0	2
382	59	1	0	110	239	0	0	142	1	1.2	1	1	3
805	53	1	0	142	226	0	0	111	1	0.0	2	0	3
145	40	1	0	110	167	0	0	114	1	2.0	1	0	3
1000	64	1	0	145	212	0	0	132	0	2.0	1	2	1
966	58	1	0	128	259	0	0	130	1	3.0	1	2	3
453	49	0	1	134	271	0	1	162	0	0.0	1	0	2
958	65	0	2	140	417	1	0	157	0	0.8	2	1	2
738	58	1	0	125	300	0	0	171	0	0.0	2	2	3
138	47	1	2	138	257	0	0	156	0	0.0	2	0	2
683	43	1	0	120	177	0	0	120	1	2.5	1	0	3
161	51	1	2	100	222	0	1	143	1	1.2	1	0	2
229	66	0	0	178	228	1	1	165	1	1.0	1	2	3
376	59	1	1	140	221	0	1	164	1	0.0	2	0	2
5	58	0	0	100	248	0	0	122	0	1.0	1	0	2
928	54	1	0	124	266	0	0	109	1	2.2	1	1	3
840	70	1	0	145	174	0	1	125	1	2.6	0	0	3
408	46	0	0	138	243	0	0	152	1	0.0	1	0	2
275	56	1	0	130	283	1	0	103	1	1.6	0	0	3
41	54	1	2	120	258	0	0	147	0	0.4	1	0	3
786	67	1	0	125	254	1	1	163	0	0.2	1	2	3
529	69	1	3	160	234	1	0	131	0	0.1	1	1	2
200	62	0	0	124	209	0	1	163	0	0.0	2	0	2
18	50	0	1	120	244	0	1	162	0	1.1	2	0	2
367	48	1	1	110	229	0	1	168	0	1.0	0	0	3
235	64	1	3	170	227	0	0	155	0	0.6	1	0	3
578	62	0	0	140	394	0	0	157	0	1.2	1	0	2
368	58	1	2	105	240	0	0	154	1	0.6	1	0	3
810	40	1	0	110	167	0	0	114	1	2.0	1	0	3
892	54	1	0	110	239	0	1	126	1	2.8	1	1	3

no data	age	sex	cp	terstbps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
736	59	1	0	110	239	0	0	142	1	1.2	1	1	3
131	51	0	2	130	256	0	0	149	0	0.5	2	0	2
620	57	1	0	152	274	0	1	88	1	1.2	1	1	3
266	53	1	0	123	282	0	1	95	1	2.0	1	2	3
987	57	1	1	154	232	0	0	164	0	0.0	2	1	2
663	58	0	0	100	248	0	0	122	0	1.0	1	0	2
35	46	1	2	150	231	0	1	147	0	3.6	1	0	2
48	66	0	2	146	278	0	0	152	0	0.0	1	1	2
858	68	0	2	120	211	0	0	115	0	1.5	1	0	2
532	63	0	0	124	197	0	1	136	1	0.0	1	0	2
512	44	1	0	112	290	0	0	153	0	0.0	2	1	2
246	54	1	1	192	283	0	0	195	0	0.0	2	1	3
727	56	1	1	130	221	0	0	163	0	0.0	2	0	3
483	35	1	1	122	192	0	1	174	0	0.0	2	0	2
955	62	1	2	130	231	0	1	146	0	1.8	1	3	3
449	52	1	0	112	230	0	1	160	0	0.0	2	1	2
102	54	1	1	108	309	0	1	156	0	0.0	2	0	3
52	38	1	2	138	175	0	1	173	0	0.0	2	4	2
775	51	1	2	100	222	0	1	143	1	1.2	1	0	2
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2
381	58	1	2	132	224	0	0	173	0	3.2	2	2	3
482	51	1	0	140	298	0	1	122	1	4.2	1	3	3
834	42	1	0	136	315	0	1	125	1	1.8	1	0	1
505	44	0	2	118	242	0	1	149	0	0.3	1	1	2
499	46	0	2	142	177	0	0	160	1	1.4	0	0	2
936	43	1	2	130	315	0	1	162	0	1.9	2	1	2
670	61	0	0	130	330	0	0	169	0	0.0	2	0	2
653	56	1	0	130	283	1	1	103	1	1.6	0	0	3
252	55	1	0	132	353	0	1	132	1	1.2	1	1	3
279	41	0	1	105	198	0	1	168	0	0.0	2	1	2
317	63	0	2	135	252	0	0	172	0	0.0	2	0	2
568	54	0	2	160	201	0	1	163	0	0.0	2	1	2
672	68	1	2	118	277	0	1	151	0	1.0	2	1	3
57	48	1	1	130	245	0	0	180	0	0.2	1	0	2
184	45	1	1	128	308	0	0	170	0	0.0	2	0	2
950	67	1	0	160	286	0	0	108	1	1.5	1	3	2
560	58	0	0	130	197	0	1	131	0	0.6	1	0	2
55	55	1	0	140	217	0	1	111	1	5.6	0	0	3
820	61	1	0	140	207	0	0	138	1	1.9	2	1	3
28	56	1	2	130	256	1	0	142	1	0.6	1	1	1
387	77	1	0	125	304	0	0	162	1	0.0	2	3	2
327	57	1	0	150	276	0	0	112	1	0.6	1	1	1
226	49	1	2	118	149	0	0	126	0	0.8	2	3	2

no data	age	sex	cp	terstbpps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
664	66	0	3	150	226	0	1	114	0	2.6	0	0	2
784	54	1	2	150	232	0	0	165	0	1.6	2	0	3
594	64	1	0	145	212	0	0	132	0	2.0	1	2	1
828	44	1	2	130	233	0	1	179	1	0.4	2	0	2
839	51	1	0	140	261	0	0	186	1	0.0	2	0	2
54	55	1	0	140	217	0	1	111	1	5.6	0	0	3
117	43	1	0	120	177	0	0	120	1	2.5	1	0	3
72	56	1	0	125	249	1	0	144	1	1.2	1	1	2
369	51	1	2	110	175	0	1	123	0	0.6	2	0	2
379	52	1	3	118	186	0	0	190	0	0.0	1	0	1
902	62	0	0	140	268	0	0	160	0	3.6	0	2	2
994	59	1	0	110	239	0	0	142	1	1.2	1	1	3
378	67	1	0	120	237	0	1	71	0	1.0	1	0	2
31	50	0	1	120	244	0	1	162	0	1.1	2	0	2
760	57	1	0	150	276	0	0	112	1	0.6	1	1	1
463	43	1	0	110	211	0	1	161	0	0.0	2	0	3
36	51	1	3	125	213	0	0	125	1	1.4	2	1	2
162	77	1	0	125	304	0	0	162	1	0.0	2	3	2
320	53	0	0	130	264	0	0	143	0	0.4	1	0	2
787	51	1	0	140	298	0	1	122	1	4.2	1	3	3
930	65	1	0	135	254	0	0	127	0	2.8	1	1	3
267	67	1	0	120	237	0	1	71	0	1.0	1	0	2
326	54	1	1	192	283	0	0	195	0	0.0	2	1	3
395	66	0	2	146	278	0	0	152	0	0.0	1	1	2
847	61	1	0	138	166	0	0	125	1	3.6	1	1	2
476	57	1	0	165	289	1	0	124	0	1.0	1	3	3
802	63	0	0	124	197	0	1	136	1	0.0	1	0	2
260	44	1	2	120	226	0	0	169	0	0.0	2	0	2
204	66	0	2	146	278	0	0	152	0	0.0	1	1	2
100	43	0	2	122	213	0	1	165	0	0.2	1	0	2
203	64	1	3	170	227	0	0	155	0	0.6	1	0	3
881	57	1	1	154	232	0	0	164	0	0.0	2	1	2
682	59	1	0	164	176	1	0	90	0	1.0	1	2	1
992	50	0	0	110	254	0	0	159	0	0.0	2	0	2
151	54	1	1	192	283	0	0	195	0	0.0	2	1	3
485	62	1	1	120	281	0	0	103	0	1.4	1	1	3
794	61	1	3	134	234	0	1	145	0	2.6	1	2	2
699	35	1	0	126	282	0	0	156	1	0.0	2	0	3
546	70	1	0	130	322	0	0	109	0	2.4	1	3	2
294	56	0	0	200	288	1	0	133	1	4.0	0	2	3
702	71	0	1	160	302	0	1	162	0	0.4	2	2	2
965	76	0	2	140	197	0	2	116	0	1.1	1	0	2
185	57	1	0	165	289	1	0	124	0	1.0	1	3	3

no data	age	sex	cp	terstbpps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
951	62	0	2	130	263	0	1	97	0	1.2	1	1	3
898	42	0	0	102	265	0	0	122	0	0.6	1	0	2
962	52	1	0	108	233	1	1	147	0	0.1	2	3	3
78	52	1	1	134	201	0	1	158	0	0.8	2	1	2
531	65	0	2	155	269	0	1	148	0	0.8	2	0	2
417	52	1	2	138	223	0	1	169	0	0.0	2	4	2
352	57	1	0	110	335	0	1	143	1	3.0	1	1	3
195	59	1	2	150	212	1	1	157	0	1.6	2	0	2
79	52	1	1	134	201	0	1	158	0	0.8	2	1	2
177	64	1	0	120	246	0	0	96	1	2.2	0	1	2
492	64	1	2	125	309	0	1	131	1	1.8	1	0	3
314	40	1	3	140	199	0	1	178	1	1.4	2	0	3
952	54	0	2	135	304	1	1	170	0	0.0	2	0	2
239	62	0	0	150	244	0	1	154	1	1.4	1	0	2
876	61	1	0	140	207	0	0	138	1	1.9	2	1	3
429	47	1	2	108	243	0	1	152	0	0.0	2	0	2
637	58	0	2	120	340	0	1	172	0	0.0	2	0	2
895	62	0	2	130	263	0	1	97	0	1.2	1	1	3
650	62	1	0	120	267	0	1	99	1	1.8	1	2	3
399	63	1	3	145	233	1	0	150	0	2.3	0	0	1
947	54	0	2	160	201	0	1	163	0	0.0	2	1	2
91	62	0	0	124	209	0	1	163	0	0.0	2	0	2
170	43	1	0	150	247	0	1	171	0	1.5	2	0	2
605	71	0	2	110	265	1	0	130	0	0.0	2	1	2
90	54	0	2	108	267	0	0	167	0	0.0	2	0	2
75	47	1	2	138	257	0	0	156	0	0.0	2	0	2
793	60	0	2	120	178	1	1	96	0	0.0	2	0	2
1003	52	1	0	108	233	1	1	147	0	0.1	2	3	3
345	66	1	1	160	246	0	1	120	1	0.0	1	3	1
927	42	1	2	120	240	1	1	194	0	0.8	0	0	3
705	58	1	0	128	216	0	0	131	1	2.2	1	3	3
734	52	1	0	128	204	1	1	156	1	1.0	1	0	0
97	53	1	0	123	282	0	1	95	1	2.0	1	2	3
931	40	1	0	152	223	0	1	181	0	0.0	2	0	3
706	57	1	2	128	229	0	0	150	0	0.4	1	1	3
914	66	1	0	112	212	0	0	132	1	0.1	2	1	2
444	47	1	2	130	253	0	1	179	0	0.0	2	0	2
968	53	1	0	140	203	1	0	155	1	3.1	0	0	3
733	44	0	2	108	141	0	1	175	0	0.6	1	0	2
1002	66	1	0	112	212	0	0	132	1	0.1	2	1	2
541	44	1	2	140	235	0	0	180	0	0.0	2	0	2
236	49	1	2	118	149	0	0	126	0	0.8	2	3	2
875	60	1	0	130	206	0	0	132	1	2.4	1	2	3

no data	age	sex	cp	terstbpps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
74	48	1	0	130	256	1	0	150	1	0.0	2	2	3
769	71	0	2	110	265	1	0	130	0	0.0	2	1	2
59	57	1	1	154	232	0	0	164	0	0.0	2	1	2
933	38	1	3	120	231	0	1	182	1	3.8	1	0	3
901	42	0	0	102	265	0	0	122	0	0.6	1	0	2
742	63	1	0	130	330	1	0	132	1	1.8	2	3	3
554	58	1	0	100	234	0	1	156	0	0.1	2	1	3
614	51	0	0	130	305	0	1	142	1	1.2	1	0	3
11	43	0	0	132	341	1	0	136	1	3.0	1	0	3
777	53	1	0	123	282	0	1	95	1	2.0	1	2	3
269	71	0	2	110	265	1	0	130	0	0.0	2	1	2
227	44	0	2	118	242	0	1	149	0	0.3	1	1	2
932	51	0	2	140	308	0	0	142	0	1.5	2	1	2
491	57	1	2	150	168	0	1	174	0	1.6	2	0	2
779	34	0	1	118	210	0	1	192	0	0.7	2	0	2
989	71	0	1	160	302	0	1	162	0	0.4	2	2	2
961	58	0	0	100	248	0	0	122	0	1.0	1	0	2
799	43	1	0	115	303	0	1	181	0	1.2	1	0	2
741	41	0	2	112	268	0	0	172	1	0.0	2	0	2
310	61	1	0	120	260	0	1	140	1	3.6	1	1	3
291	58	1	0	128	259	0	0	130	1	3.0	1	2	3
856	68	0	2	120	211	0	0	115	0	1.5	1	0	2
703	57	1	0	110	201	0	1	126	1	1.5	1	0	1
308	59	1	2	126	218	1	1	134	0	2.2	1	1	1
299	52	1	1	120	325	0	1	172	0	0.2	2	0	2
597	38	1	2	138	175	0	1	173	0	0.0	2	4	2
363	53	1	2	130	246	1	0	173	0	0.0	2	3	2
719	52	1	0	108	233	1	0	147	0	0.1	2	3	3
431	65	0	0	150	225	0	0	114	0	1.0	1	3	3
713	66	0	3	150	226	0	1	114	0	2.6	0	0	2
107	62	1	1	120	281	0	0	103	0	1.4	1	1	3
542	62	0	0	140	394	0	0	157	0	1.2	1	0	2
829	54	1	0	124	266	0	0	109	1	2.2	1	1	3
584	55	1	0	132	353	0	1	132	1	1.2	1	1	3
690	58	1	2	105	240	0	0	154	1	0.6	1	0	3
822	60	1	0	130	253	0	1	144	1	1.4	2	1	3
421	42	0	0	102	265	0	0	122	0	0.6	1	0	2
406	58	1	2	140	211	1	0	165	0	0.0	2	0	2
313	74	0	1	120	269	0	0	121	1	0.2	2	1	2
351	50	1	0	144	200	0	0	126	1	0.9	1	0	3
925	57	1	0	130	131	0	1	115	1	1.2	1	1	3
831	58	1	1	125	220	0	1	144	0	0.4	1	4	3
696	50	0	2	120	219	0	1	158	0	1.6	1	0	2

no data	age	sex	cp	terstbps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
373	58	1	1	120	284	0	0	160	0	1.8	1	0	2
991	60	1	0	117	230	1	1	160	1	1.4	2	2	3
328	70	1	0	130	322	0	0	109	0	2.4	1	3	2
343	52	1	2	172	199	1	1	162	0	0.5	2	0	3
660	61	1	0	138	166	0	0	125	1	3.6	1	1	2
237	57	1	1	124	261	0	1	141	0	0.3	2	0	3
461	48	1	2	124	255	1	1	175	0	0.0	2	2	2
316	64	0	2	140	313	0	1	133	0	0.2	2	0	3
954	53	0	0	130	264	0	0	143	0	0.4	1	0	2
948	70	1	0	145	174	0	1	125	1	2.6	0	0	3
921	58	1	0	146	218	0	1	105	0	2.0	1	1	3
893	52	1	0	128	204	1	1	156	1	1.0	1	0	0
748	44	1	2	120	226	0	1	169	0	0.0	2	0	2
240	54	0	1	132	288	1	0	159	1	0.0	2	1	2
972	52	1	3	118	186	0	0	190	0	0.0	1	0	1
619	65	1	0	110	248	0	0	158	0	0.6	2	2	1
586	64	1	2	125	309	0	1	131	1	1.8	1	0	3
607	42	1	0	136	315	0	1	125	1	1.8	1	0	1
178	44	1	0	110	197	0	0	177	0	0.0	2	1	2
1023	50	0	0	110	254	0	0	159	0	0.0	2	0	2
243	60	1	2	140	185	0	0	155	0	3.0	1	0	2
844	60	1	0	140	293	0	0	170	0	1.2	1	2	3
202	52	1	3	152	298	1	1	178	0	1.2	1	0	3
300	46	0	1	105	204	0	1	172	0	0.0	2	0	2
590	74	0	1	120	269	0	0	121	1	0.2	2	1	2
230	58	1	0	125	300	0	0	171	0	0.0	2	2	3
306	44	0	2	118	242	0	1	149	0	0.3	1	1	2
166	56	1	0	132	184	0	0	105	1	2.1	1	1	1
656	57	0	1	130	236	0	0	174	0	0.0	1	1	2
704	51	1	2	94	227	0	1	154	1	0.0	2	1	3
263	55	0	1	132	342	0	1	166	0	1.2	2	0	2
346	50	1	0	150	243	0	0	128	0	2.6	1	0	3
232	60	1	0	125	258	0	0	141	1	2.8	1	1	3
150	58	1	0	114	318	0	2	140	0	4.4	0	3	1
498	54	1	0	140	239	0	1	160	0	1.2	2	0	2
132	41	1	1	135	203	0	1	132	0	0.0	1	0	1
9	54	1	0	122	286	0	0	116	1	3.2	1	2	2
254	35	1	0	120	198	0	1	130	1	1.6	1	0	3
891	64	0	0	180	325	0	1	154	1	0.0	2	0	2
561	65	0	2	155	269	0	1	148	0	0.8	2	0	2
669	66	1	1	160	246	0	1	120	1	0.0	1	3	1
82	46	1	2	150	231	0	1	147	0	3.6	1	0	2
64	29	1	1	130	204	0	0	202	0	0.0	2	0	2

no data	age	sex	cp	terstbps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
633	61	1	0	138	166	0	0	125	1	3.6	1	1	2
853	67	1	0	120	229	0	0	129	1	2.6	1	2	3
613	55	1	0	140	217	0	1	111	1	5.6	0	0	3
548	51	0	2	120	295	0	0	157	0	0.6	2	0	2
478	39	1	2	140	321	0	0	182	0	0.0	2	0	2
119	42	1	1	120	295	0	1	162	0	0.0	2	0	2
791	54	1	0	110	239	0	1	126	1	2.8	1	1	3
729	55	0	1	135	250	0	0	161	0	1.4	1	0	2
689	54	0	2	135	304	1	1	170	0	0.0	2	0	2
538	46	1	0	120	249	0	0	144	0	0.8	2	0	3
110	44	1	0	110	197	0	0	177	0	0.0	2	1	2
692	53	1	0	140	203	1	0	155	1	3.1	0	0	3
918	59	1	3	134	204	0	1	162	0	0.8	2	2	2
937	64	1	2	140	335	0	1	158	0	0.0	2	0	2
437	67	1	0	120	229	0	0	129	1	2.6	1	2	3
121	44	1	0	120	169	0	1	144	1	2.8	0	0	1
632	45	0	1	130	234	0	0	175	0	0.6	1	0	2
789	62	1	1	120	281	0	0	103	0	1.4	1	1	3
797	65	0	0	150	225	0	0	114	0	1.0	1	3	3
833	55	1	0	140	217	0	1	111	1	5.6	0	0	3
521	58	1	1	125	220	0	1	144	0	0.4	1	4	3
698	66	1	0	112	212	0	0	132	1	0.1	2	1	2
524	58	1	2	112	230	0	0	165	0	2.5	1	1	3
911	58	0	1	136	319	1	0	152	0	0.0	2	2	2
744	60	0	2	102	318	0	1	160	0	0.0	2	1	2
40	65	0	2	160	360	0	0	151	0	0.8	2	0	2
756	64	1	0	120	246	0	0	96	1	2.2	0	1	2
735	50	1	2	129	196	0	0	163	0	0.0	2	0	2
217	41	1	1	120	157	0	1	182	0	0.0	2	0	2
307	44	1	1	120	220	0	1	170	0	0.0	2	0	2
543	59	1	3	134	204	0	1	162	0	0.8	2	2	2
850	58	1	1	120	284	0	0	160	0	1.8	1	0	2
157	54	1	2	120	258	0	0	147	0	0.4	1	0	3
250	47	1	0	110	275	0	0	118	1	1.0	1	1	2
124	61	1	0	120	260	0	1	140	1	3.6	1	1	3
743	58	1	1	125	220	0	1	144	0	0.4	1	4	3
329	53	0	2	128	216	0	0	115	0	0.0	2	0	0
349	62	0	2	130	263	0	1	97	0	1.2	1	1	3
528	59	1	3	178	270	0	0	145	0	4.2	0	0	3
384	35	1	0	126	282	0	0	156	1	0.0	2	0	3
169	45	0	1	112	160	0	1	138	0	0.0	1	0	2
397	40	1	0	110	167	0	0	114	1	2.0	1	0	3
714	51	1	3	125	213	0	0	125	1	1.4	2	1	2

no data	age	sex	cp	terstbps	chol	fbs	restecg	thalac	exang	oldpeak	slope	ca	thal
581	44	1	2	140	235	0	0	180	0	0.0	2	0	2
34	50	1	2	129	196	0	1	163	0	0.0	2	0	2
807	44	1	2	130	233	0	1	179	1	0.4	2	0	2
309	56	0	1	140	294	0	0	153	0	1.3	1	0	2
788	62	0	0	138	294	1	1	106	0	1.9	1	3	2
163	48	1	0	124	274	0	0	166	0	0.5	1	0	3
370	43	0	0	132	341	1	0	136	1	3.0	1	0	3
159	41	1	1	120	157	0	1	182	0	0.0	2	0	2
287	71	0	1	160	302	0	1	162	0	0.4	2	2	2
183	42	1	2	130	180	0	1	150	0	0.0	2	0	2
626	58	1	2	132	224	0	0	173	0	3.2	2	2	3
879	59	1	0	138	271	0	0	182	0	0.0	2	0	2
587	59	1	0	164	176	1	0	90	0	1.0	1	2	1
21	67	0	0	106	223	0	1	142	0	0.3	2	2	2
761	49	0	0	130	269	0	1	163	0	0.0	2	0	2
983	64	1	0	128	263	0	1	105	1	0.2	1	1	3
283	64	1	3	110	211	0	0	144	1	1.8	1	0	2
479	58	1	0	128	216	0	0	131	1	2.2	1	3	3
1004	51	0	2	140	308	0	0	142	0	1.5	2	1	2
746	64	1	2	140	335	0	1	158	0	0.0	2	0	2
595	61	1	0	148	203	0	1	161	0	0.0	2	1	3
160	77	1	0	125	304	0	0	162	1	0.0	2	3	2
976	63	0	0	108	269	0	1	169	1	1.8	1	2	2
120	54	1	1	108	309	0	1	156	0	0.0	2	0	3

