



## Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi Menggunakan Algoritma C5.0

Dito Putro Utomo<sup>1,2,\*</sup>, Pahala Sirait<sup>1</sup>, Roni Yunis<sup>1</sup>

<sup>1</sup> Program Studi Magister Teknologi Informasi, STMIK Mikroskil, Medan, Indonesia

<sup>2</sup> Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email Penulis Korespondensi: ditoputro12@gmail.com

**Abstrak**—Penyakit jantung koroner biasa disebut dengan kardiovaskular, penyakit jantung merupakan penyakit dengan tingkat kematian yang tinggi. Dengan demikian diagnosa sangat penting dan merupakan bidang penting dari penelitian medis. Dalam proses diagnosa hal yang paling sering dihadapi adalah waktu di dalam pengambilan keputusan dan kurangnya akurasi pada proses klasifikasi. Atribut merupakan hal penting di dalam pengambilan keputusan pada penyakit jantung sehingga perlu diketahui atribut utama pada penyakit jantung, seringkali didapatkan hasil yang berbeda pada proses diagnosa dikarenakan banyaknya atribut yang digunakan pada pengambilan keputusan. Sehingga perlu dilakukannya proses reduksi pada atribut penyakit jantung. Metode Principal Componen Analysis (PCA) dapat digunakan untuk reduksi data yang berdimensi besar serta melakukan perangkingan terhadap atribut yang akan direduksi. Proses klasifikasi dapat dilakukan dengan menggunakan Algoritma C5.0 dan mendapatkan tingkat akurasi pada proses klasifikasi. Hasil yang didapatkan pada penelitian mereduksi 12 atribut pada dataset penyakit jantung dan melakukan klasifikasi dengan kombinasi atribut setelah dilakukan proses reduksi. Hasil yang didapatkan dengan tingkat akurasi tertinggi ketika dilakukan klasifikasi dengan 11 kombinasi atribut dimana terdapat 1 atribut yang direduksi, tingkat akurasi yang didapatkan sebesar 89,11%.

**Kata Kunci:** Data Mining, Reduksi, Penyakit Jantung, Principal Compnent Analysis, C5.0.

**Abstract**—Coronary heart disease, commonly referred to as cardiovascular, heart disease is a disease with a high mortality rate. Thus diagnosis is very important and is an important area of medical research. In the diagnostic process, the most frequently encountered problems are time in making decisions and the lack of accuracy in the classification process. Attributes are important in making decisions on heart disease so it is necessary to know the main attributes of heart disease. Often different results are obtained in the diagnostic process due to the many attributes used in decision making. So it is necessary to do a reduction process in the attributes of heart disease. Principal Component Analysis (PCA) method can be used for data reduction with large dimensions and ranking the attributes to be reduced. The classification process can be done using the C5.0 Algorithm and getting a level of accuracy in the classification process. The results obtained in this study reduce the 12 attributes of the heart disease dataset and classify them with a combination of attributes after the reduction process is carried out. The results obtained with the highest level of accuracy when classifying with 11 attribute combinations where there is 1 attribute that is reduced, the accuracy rate obtained is 89.11%.

**Keywords:** Data Mining, Reduction, Heart Disease, Principal Compnent Analysis, C5.0.

### 1. PENDAHULUAN

Penyakit jantung koroner biasa juga disebut dengan kardiovaskular, yaitu menyempit atau tersumbat pembuluh darah yang mengarah pada serangan jantung, nyeri dada atau stroke [1]. Penyakit jantung merupakan penyakit dengan tingkat kematian yang tinggi dengan setiap tahun lebih dari 12 juta kematian terjadi diseluruh dunia dikarenakan penyakit jantung [2]. Dengan demikian diagnosis dini sangat penting dan merupakan bidang penting dari penelitian medis, [3]. Diagnosis yang benar dari penyakit jantung adalah tugas yang cukup menantang karena saling ketergantungan yang kompleks dari berbagai faktor [4]. Permasalahan yang sering kali dihadapi adalah waktu di dalam proses serta kurangnya keakurasian pada proses klasifikasi [5].

Fitur atribut merupakan suatu hal yang penting di dalam keakuratan hasil, sehingga perlu kiranya diketahui atribut utama pada sebuah penyakit [6]. Sering didapatkan hasil yang berbeda ketika mendiagnosis penyakit jantung, sehingga dipilih fitur atribut penting yang didasarkan pada efisiensi diagnosis [7]. Keputusan diagnosis penyakit yang berbeda merupakan salah satu tantangan yang paling penting untuk analisis data penyakit jantung [1]. Umumnya keputusan diagnosis didasarkan kepada pemeriksaan kepada pasien dan juga pengalaman dokter. Pada pemeriksaan yang dilakukan tidak semua kondisi atribut terpenuhi, tetapi tetap harus dilakukan proses diagnosis.

Kesalahan pada proses diagnosis akan mempengaruhi hasil pada pasien ataupun dokter, jadi diagnosis merupakan tugas kompleks yang membutuhkan keahlian tinggi dan pengalaman. Sistem komputer dapat membantu para dokter sebagai alat memprediksi dan diagnosis penyakit jantung, penelitian medis selalu berurusan dengan data yang besar, penanganan data yang besar dengan tidak benar dapat mempengaruhi keakuratan terhadap hasil. Dataset yang digunakan merupakan kumpulan dari pengalaman dokter dan juga dari beberapa rumah sakit serta perguruan tinggi [2]. Reduksi atribut merupakan pemilihan fitur digunakan untuk mengurangi dimensi data dengan menghapus atribut yang berlebihan dan tidak relevan dalam satu set data. Proses klasifikasi dipengaruhi oleh atribut, fitur redundansi yang tidak relevan memiliki efek negatif pada keakuratan klasifikasi dan setelah dilakukan redundansi pada atribut, tidak mengurangi tingkat akurasi terhadap dataset. [8]. Pengurangan atribut telah menjadi langkah penting dalam pengenalan pola dan tugas Machine Learning [9]. Hal ini menginspirasi



penulis sehingga menambahkan proses reduksi atribut di dalam pengambilan keputusan klasifikasi pada dataset penyakit jantung.

Principal component analysis (PCA) adalah teknik untuk mengurangi dimensi atribut dari dataset, meningkatkan interpretabilitas tetapi pada saat yang sama meminimalkan hilangnya informasi, banyak penelitian dan buku yang telah ditulis dan juga bahkan ada buku tentang varian PCA untuk tipe data tertentu [10]. PCA dapat digabungkan dengan metode lain yang akurasi dikenal keakurasiannya cukup tinggi, dimana PCA berfungsi untuk melakukan reduksi dimensionalitas dan dengan pemilihan fitur pada PCA yang hanya memilih nilai eigen yang paling besar maka ketika digabungkan dengan metode lain akan menaikkan tingkat akurasi klasifikasi dan terutama sekali mempercepat hasil pengenalannya [11].

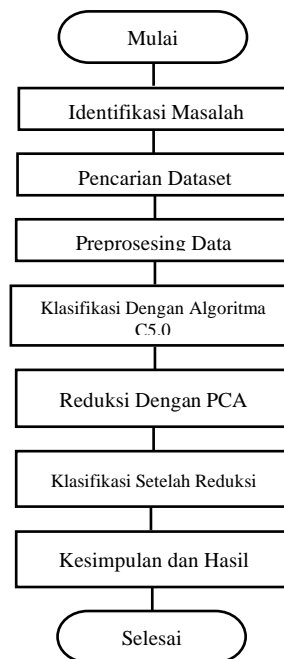
Algoritma C5.0 memiliki kemampuan untuk mengklasifikasikan sebagai pohon keputusan atau seperangkat aturan. Kekuatan algoritma ini dalam menangani nilai, serta lebih sedikit waktu diperlukan untuk mempelajarinya [12]. C5.0 adalah classifier yang mengklasifikasikan data dalam waktu yang lebih singkat dibandingkan dengan classifier lain. Untuk menghasilkan pohon keputusan penggunaan memori minimum dan juga meningkatkan akurasi [13]. C5.0 merupakan terusan dari algoritma C4.5, dimana C5.0 memperbaiki kelemahan dari C4.5 dari segi hasil klasifikasi, kesalahan klasifikasi, prediksi akurasi, waktu, tingkat kesalahan, dan pemakaian memori [14]

Berdasarkan dari penjelasan di atas maka penulis bermaksud untuk melakukan penelitian sebagai berikut, melakukan reduksi atribut pada dataset penyakit jantung. Setelah dilakukan reduksi atribut kemudian dilakukan klasifikasi kembali terhadap atribut-atribut yang digunakan selanjutnya dan apakah terdapat perubahan pada tingkat akurasi klasifikasi dari proses yang setelah terdapat reduksi atribut dengan proses klasifikasi sebelum diadakannya reduksi atribut. Pada proses reduksi atribut dan juga klasifikasi penelitian ini menggunakan dataset yang dibuat oleh para ahli dibidang penyakit jantung. Tetapi sebelum dilakukannya proses reduksi atribut dan juga klasifikasi, dilakukan terlebih dahulu preprosesing pada dataset. Dimana tujuan dilakukan preprosesing untuk menyamakan karakteristik dataset terhadap metode yang akan digunakan.

## 2. METODOLOGI PENELITIAN

### 2.1 Metodologi Penelitian

Metodologi penelitian merupakan tahapan-tahapan yang sistematis dilakukan pada penelitian ini sehingga penelitian ini terarah dengan baik. Berikut adalah metodologi penelitian yang dilakukan



Gambar 1. Metodologi Penelitian

### 2.2 Penelitian Terkait

Penelitian yang dilakukan oleh [6] memprediksi pasien yang terkena penyakit jantung dengan menggunakan fitur memiliki korelasi tertinggi dengan penyakit jantung, fitur yang berbeda diukur dan juga dikumpulkan dari pasien dan diproses.

Pada penelitian yang dilakukan oleh [2] mengatakan sistem komputer dapat membantu dokter untuk memprediksi dan mendiagnosis penyakit jantung yang berurusan dengan sejumlah data yang besar dan teratur,



penanganan data secara tradisional dapat mempengaruhi hasil klasifikasi. Sehingga diperlukan sebuah teknik dan alat untuk mengetahui informasi-informasi yang tersimpan pada database.

Penelitian lainnya yang telah dilakukan oleh [1] memprediksi dan mencari faktor resiko paling penting pada penyakit jantung dengan menggunakan dataset dari UCI Repository Machine Learning.

Penelitian yang dilakukan oleh [10] mengurangi dimensionalitas pada dataset, meningkatkan interpretabilitas tetapi meminimalkan hilangnya informasi menggunakan metode PCA, dengan membuat variable tidak berkorelasi dan memaksimalkan varians sehingga menghasilkan variabel komponen baru.

Pada penelitian lainnya yang dilakukan oleh [16] melakukan pengurangan dimensi atribut pada data set KDD Cup dan UNB ISCX pada deteksi intrusi jaringan, berfokus pada pengurangan rasio dan hasil yang didapatkan dimensi awal pada data set sebanyak 41 atribut kemudian direduksi hingga terdapat 10 atribut utama dengan tingkat akurasi sebesar 99,7% dan 98,8%.

Penelitian lain yang telah dilakukan oleh [17] mendeteksi penyakit Multiple Sclerosis (MS) menggunakan 3 metode Biorthogonal Wavelet Transform, Principal Component Analysis, Logistic Regression. Hasil analisis mencapai sensitivitas  $97,12 \pm 0,14\%$ , spesifisitas dari  $98,25 \pm 0,16\%$ , dan akurasi  $97,76 \pm 0,10\%$ .

Penelitian lain yang dilakukan oleh [18] melakukan reduksi pada citra medis dengan menghilangkan noise, mengklasifikasikan dan mengekstrak fitur citra medis. PCA digunakan sebagai kompresi citra dengan tingkat kehilangan informasi yang rendah dan menunjukkan efisiensi yang tinggi.

Penelitian yang dilakukan oleh [12] melakukan prediksi kemungkinan pasien menderita penyakit jantung dengan menggunakan 5 metode yaitu : C5.0, Neural Network, SVM, KNN, dan Logistic Regression. Hasil yang didapatkan adalah pohon keputusan C5.0 memiliki tingkat akurasi yang paling tinggi dibandingkan dengan metode-metode yang lainnya.

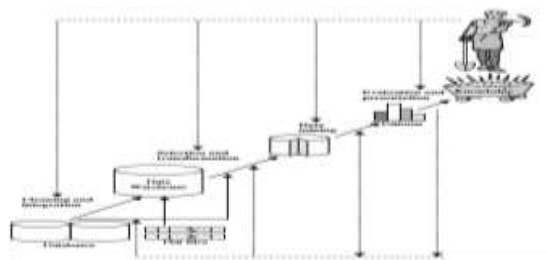
Penelitian lainnya yang dilakukan oleh [19] melakukan prediksi pada penyakit hati manusia, dengan menggunakan data set dari UCI Repository Machine Learning. Proses klasifikasi dan juga prediksi menggunakan algoritma C5.0 dan CHAID, hasil yang didapatkan algoritma C5.0 mendapatkan akurasi yang lebih baik sebesar 93.75% dibandingkan dengan CHAID sebesar 65%. Serta mendapatkan faktor-faktor yang paling signifikan di dalam memprediksi penyakit hati.

Penelitian lainnya yang dilakukan oleh [14] melakukan perbandingan diantara algoritma C4.5 dan C5.0, dimana C5.0 merupakan perbaikan dari C4.5 dari segi dari hasil klasifikasi, kesalahan klasifikasi, prediksi akurasi, waktu, tingkat kesalahan, dan pemakaian memori.

## 2.3 Data Mining

Data mining adalah proses menemukan data yang tersedia guna membuat sebuah model baru, kemudian menggunakan model tersebut untuk memahami pola data lain yang tidak tersedia didalam database yang tersedia. Pada data mining proses penentuan kelompok data bisa juga dilaksanakan guna menemukan pola secara keseluruhan pada data yang tersedia untuk dilakukan proses tindak lanjut lainnya yang berguna sebagai pendukung kegiatan dan tujuan akhir tertentu. Tujuan pada data mining untuk menemukan hubungan ataupun pola yang memiliki kemungkinan menyajikan informasi yang bermanfaat. Dari pengertian tersebut bisa diambil sebuah kesimpulan bahwa data mining adalah proses menemukan serta pengenalan pola dan juga proses pengelompokan data-data yang besar serta bertujuan guna mendapatkan hubungan atau pola yang memberikan manfaat.

Pada data mining terdapat istilah yang memiliki makna sama dengan data mining yaitu Knowledge Discovery in Database (KDD). Data mining dan KDD mempunyai tujuan yang sama yaitu memanfaatkan data yang tersedia pada database lalu mengolah data guna menghasilkan informasi baru yang bermanfaat. Selain itu, terdapat banyak istilah yang memiliki makna serupa dengan data mining misalnya, Knowledge Discovery in Database, ekstraksi pengetahuan, analisis pola / data, arkeologi data, dan pengerukan data. Banyak orang memperlakukan data mining sebagai sinonim untuk istilah lain yang populer digunakan, penemuan pengetahuan dari data, atau KDD, sementara yang lain melihat penambahan data hanya sebagai langkah penting dalam proses penemuan pengetahuan [15].



Gambar 2. Tahapan Proses Knowledge Discovery in Database (KDD)

## 2.4 Metode Principal Component Analysis (PCA)

Principal Component Analysis (PCA) melakukan pemetaan data dari dimensi lama ke dimensi baru dengan menggunakan teknik aljabar linear, tanpa harus menggunakan parameter tertentu untuk menghasilkan keluaran



metanya. PCA menggunakan input data yang memiliki sifat zero-mean pada setiap atributnya. Sifat zero-mean pada atribut data didapatkan dengan melakukan mengurangkan semua nilai dengan rata-ratanya. Dataset X dengan dimensi  $M \times N$ , dimana M adalah jumlah data dan N adalah jumlah fitur.

$$X = \begin{bmatrix} X_{11} & X_{12} & X_{1j} & X_{1N} \\ X_{21} & \dots & \dots & X_{2N} \\ X_{31} & \dots & \dots & X_{3N} \\ \dots & \dots & \dots & \dots \\ X_{i1} & \dots & \dots & X_{iN} \\ \dots & \dots & \dots & \dots \\ X_{M1} & X_{M2} & X_{Mj} & X_{MN} \end{bmatrix}$$

Untuk fitur ke-j, semua nilai pada kolom tersebut dikurangi dengan rata-ratanya, diformulasikan dengan:

$$X_{ij} = X_{ij} - \bar{X}_j \tag{1}$$

$i = 1, 2, \dots, M$ , dan  $j$  adalah kolom ke-j

Selanjutnya dilakukan perhitungan matriks kovarian dari matriks X, yaitu  $C_X$ . Formula yang digunakan adalah dot-product pada setiap fitur.

$$C_X = \frac{1}{M} X^T \cdot X \tag{2}$$

N adalah jumlah fitur, sedangkan  $X^T$  adalah matriks transpos dari X.

$$C_X = \frac{1}{M} \begin{bmatrix} X_{11} & X_{12} & X_{1j} & X_{1N} \\ X_{21} & \dots & \dots & X_{2N} \\ X_{i1} & \dots & \dots & X_{iN} \\ X_{M1} & X_{M2} & X_{Mj} & X_{MN} \end{bmatrix} \times \begin{bmatrix} X_{11} & X_{12} & X_{1j} & X_{1N} \\ X_{21} & \dots & \dots & X_{2N} \\ X_{i1} & \dots & \dots & X_{iN} \\ X_{M1} & X_{M2} & X_{Mj} & X_{MN} \end{bmatrix}$$

$$= \frac{1}{M} \begin{bmatrix} X_{11} & X_{12} & X_{1j} & X_{1N} \\ X_{21} & \dots & \dots & X_{2N} \\ X_{i1} & \dots & \dots & X_{iN} \\ X_{N1} & X_{N2} & X_{Nj} & X_{NN} \end{bmatrix}$$

Pada matriks  $C_X$ , elemen ke-ij adalah inner-product antara baris matriks  $X^T$  dengan kolom matriks X. Sifat-sifat yang dimiliki oleh matriks  $C_X$

Cara yang umum digunakan untuk mendapatkan  $C_Y$  adalah dengan eigenvalue dan eigenvector. Eigenvalue dan eigenvector dari matriks X berturut-turut adalah nilai skala  $\lambda$  dan vektor u yang memenuhi persamaan berikut

$$Xu = \lambda u \tag{3}$$

Dengan mencari matriks ortonormal P dimana  $Y = PX$  dan  $C_Y = \frac{1}{M} YY^T$  adalah matriks diagonal, dan kolom dari P adalah komponen utama (principal components) dari X, persamaan  $C_Y$  bisa dijabarkan sebagai berikut:

$$C_Y = \frac{1}{M} YY^T$$

$$= \frac{1}{M} (PX)(PX)^T$$

$$= \frac{1}{M} PXX^T P^T$$

$$= P \left( \frac{1}{M} XX^T \right) P^T$$

Dengan mendistribusikan persamaan 2, didapatkan matriks  $C_Y$  berdimensi  $N \times N$ :

$$C_Y = PC_X P^T \tag{4}$$

### 2.5 Teknik Klasifikasi

Klasifikasi digunakan untuk menilai objek data lalu memasukkannya pada kelas tertentu dari seluruh kelas yang tersedia. Klasifikasi melakukan proses pembangunan model didasarkan dengan data latih yang tersedia, lalu memanfaatkan model tersebut guna mengklasifikasikan terhadap data yang baru. Klasifikasi merupakan pekerjaan yang melakukan pembelajaran dari fungsi target yang mengelompokkan setiap set atribut ke label kelas yang tersedia. Sistem yang melakukan proses klasifikasi diharapkan mampu melakukan klasifikasi data set dengan benar, tetapi kinerja sistem tidak bisa 100% benar sehingga sebuah sistem klasifikasi juga harus diukur kinerjanya. Umumnya, pengukuran kinerja klasifikasi dilakukan dengan matriks konfusi.

**Tabel 1.** Matriks Konfusi Untuk Dua Kelas

$f_{ij}$	Kelas Hasil Prediksi ( $j$ )		
	Kelas = 1	Kelas = 0	
Kelas Asli ( $i$ )	Kelas = 1	$F_{11}$	$F_{10}$
	Kelas = 0	$F_{01}$	$F_{00}$

Berdasarkan matriks konfusi, bisa diketahui hasil klasifikasi dari setiap kelas yang telah diprediksi secara benar dan data yang diklasifikasikan secara salah. Kuantitasi matriks konfusi dapat diringkas menjadi dua nilai, yaitu akurasi dan laju eror. Dengan mengetahui jumlah data yang diklasifikasi secara benar, dapat mengetahui



akurasi hasil klasifikasi, dan dengan mengetahui jumlah data yang diklasifikasikan secara salah, dapat mengetahui laju eror yang dilakukan

Untuk mengetahui akurasi digunakan formula :

$$\text{Akurasi} = \frac{\text{Jumlah Data Yang Diprediksi Benar}}{\text{Jumlah Prediksi Yang Dilakukan}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

Untuk mengetahui laju eror digunakan formula :

$$\text{Akurasi} = \frac{\text{Jumlah Data Yang Diprediksi Salah}}{\text{Jumlah Prediksi Yang Dilakukan}} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

## 2.6 Algoritma C5.0

Algoritma C5.0 adalah perbaikan dari algoritma C4.5 yang juga perbaikan dari algoritma ID3. C5.0 adalah algoritma klasifikasi yang dapat dilakukan dalam kumpulan data besar. Algoritma C5.0 ini lebih baik daripada C4.5 pada kecepatan proses, memori dan efisiensi. Algoritma C5.0 bekerja dengan membedakan sampel berdasarkan pada atribut yang menyediakan informasi. C5.0 dapat membagi atribut berdasarkan dari nilai informasi gain yang paling besar. Proses akan berlanjut hingga bagian sampel tidak dapat dibagi. Algoritma C5.0 dapat menangani atribut kontinyu dan diskrit. Pemilihan atribut dalam algoritma ini akan diproses menggunakan information gain. Atribut dengan nilai Gain tertinggi akan dipilih sebagai akar bagi node selanjutnya. Berikut persamaan untuk menghitung entropy atribut[20] :

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m p_i * \log_2(p_i) \tag{5}$$

Dengan :

S : Himpunan kasus

m : Jumlah sampel

$p_i$  : Proporsi kelas

Sementara untuk mendapatkan informasi nilai subset tersebut dapat dilihat pada persamaan sebagai berikut:

$$E(A) = \sum_{j=i}^y \frac{S_{ij} + \dots + S_{mj}}{s} I(S_{1j}, \dots, S_{mj}) \tag{6}$$

Dengan :

$$\frac{S_{ij} + \dots + S_{mj}}{s} = \text{Jumlah subset J yang dibagi dengan jumlah sampel S}$$

Maka untuk mendapatkan nilai gain, selanjutnya digunakan formula. Maka nilai gain dapat dihitung dengan formula sebagai berikut:

$$\text{Gain}(A) = I(S_1, S_2, \dots, S_m) - E(A) \tag{7}$$

Dengan :

A = Atribut

S = Himpunan Kasus

$S_1$  = Jumlah Sampel

## 3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan skunder / dataset yang berasal dari UCI Repository Machine Learning, dimana terdapat 12 atribut dan 2 target kelas pada dataset penyakit jantung tersebut dan memiliki jumlah record sebanyak 303. Dataset sudah terlebih dahulu dikumpulkan oleh orang lain dan juga dataset telah digunakan pada penelitian medis terkait penyakit jantung sebelumnya

Pada Penelitian ini melakukan reduksi pada atribut dataset penyakit jantung menggunakan metode Principal Component Analysis (PCA) kemudian melakukan klasifikasi pada setiap proses reduksi atributnya hingga menemukan kombinasi atribut optimal dari dataset. Kombinasi atribut optimal dilihat dari nilai Eigenvalue dari proses reduksi dengan PCA dan tingkat akurasi dari proses klasifikasi menggunakan algoritma C5.0. Pada dataset terdapat 12 atribut antara lain : *Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slop, Ca* dan *Thal*. Selain atribut pada dataset juga memiliki 2 target kelas yaitu: *CAD* dan *Normal*.

### 3.1 Hasil Pengujian

Pada tahapan awal pengujian, terlebih dahulu melakukan klasifikasi pada dataset menggunakan algoritma C5.0. Setelah dilakukan klasifikasi lalu dilakukan proses reduksi atribut menggunakan metode Principal Component Analysis (PCA). Kombinasi reduksi atribut pada PCA dapat dilihat dari nilai eigenvalue yang didapatkan, dimana kombinasi atribut yang memiliki nilai eigenvalue paling besar akan digunakan sebagai atribut untuk proses klasifikasi selanjutnya. Adapun hasil dari proses klasifikasi dengan algoritma C5.0 sebelum dilakukan reduksi atributnya dapat dilihat pada tabel dibawah ini :

**Tabel 2.** Tingkat Akurasi Tanpa Reduksi Atribut

	True CAD	True Normal	Class Precision
Pred. CAD	162	35	82,23%



	True CAD	True Normal	Class Precision
<b>Pred. Normal</b>	3	103	97,17%
<b>Class Recall</b>	98,18%	74,64%	<b>87,46%</b>

Pada tabel 2. terdapat tingkat akurasi pada kategori yang berbeda – beda, yaitu Precision, Recall dan Accuray. Precision merupakan ketepatan tingkat akurasi antara informasi yang tersedia pada dataset dengan jawaban yang diberikan oleh hasil klasifikasi. Recall adalah tingkat keberhasilan dalam menemukan kembali sebuah informasi. Accuracy adalah tingkat kedekatan antara nilai prediksi dan nilai actual. Dari hasil klasifikasi tanpa dilakukan reduksi atribut didapatkan tingkat akurasi sebesar **87,46%**.

Kemudian setelah dilakukan klasifikasi dan mengukur tingkat akurasi pada dataset yang tidak terdapat reduksi atribut, selanjutnya melakukan reduksi 1 atribut dan mencari kombinasi 11 atribut berdasarkan nilai Eigenvalue untuk melakukan perangkingan kombinasi atribut. Adapun hasil metode Principal Component Analysis (PCA) kombinasi 11 atribut dapat dilihat di bawah ini :

**Tabel 3.** Metode PCA Kombinasi 11 Atribut

Eigenvalue	Kombinasi	Reduksi
0,7577	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal
0,7625	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca
0,7760	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope
0,7829	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak
0,7726	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang
<b>0,7832</b>	<b>Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal</b>	<b>Thalach</b>
0,7527	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg
0,7509	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	FBS
0,7527	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol
0,7558	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Trestbps
0,7624	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	CP
0,7679	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Age

Tabel 3. merupakan hasil dari proses metode Principal Componen Analysis dengan reduksi 1 atribut dan kombinasi 11 atribut. Dari hasil yang didapatkan kombinasi atribut yang memiliki nilai perangkingan terbesar berdasarkan nilai Eigenvalue adalah atribut Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slop, Ca, Thal dengan mereduksi atribut Thalach. Kemudian melakukan klasifikasi kembali dengan menggunakan kombinasi 11 atribut tersebut, sehingga didapatkan hasil klasifikasi dan tingkat akurasinya pada di bawah ini :

**Tabel 4.** Tingkat Akurasi Kombinasi 11 Atribut

	True CAD	True Normal	Class Precision
<b>Pred. CAD</b>	160	28	85,11%
<b>Pred. Normal</b>	5	110	95,65%
<b>Class Recall</b>	96,97%	79,71%	<b>89,11%</b>

Dari hasil pada tabel 4. terdapat tingkat persentase yang berbeda seperti Precision, Recall, dan Accuracy. Dari hasil klasifikasi kombinasi 11 atribut didapatkan tingkat akurasi sebesar **89,11%**. Selanjutnya setelah melakukan klasifikasi dan mengukur tingkat akurasi pada dataset dengan mereduksi 1 atribut dan kombinasi 11 atribut, melakukan kembali proses reduksi 2 atribut dan mencari kombinasi 10 atribut berdasarkan nilai Eigenvalue untuk melakukan perangkingan kombinasi atribut. Adapun hasil metode Principal Component Analysis (PCA) kombinasi 10 atribut dapat dilihat di bawah ini :

**Tabel 5.** Hasil Metode PCA Kombinasi 10 Atribut

Eigenvalue	Kombinasi	Reduksi
0,7466	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca
0,7630	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope
0,7688	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak
0,7568	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang
0,7713	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach
0,7365	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg
0,7346	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, FBS
0,7360	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Chol
0,7400	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Trestbps
0,7460	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, CP
0,7542	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Age
0,7708	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope
0,7756	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak
0,7653	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang
0,7756	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach



Eigenvalue	Kombinasi	Reduksi
0,7414	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg
0,7392	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, FBS
0,7414	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Chol
0,7446	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Trestbps
0,7518	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, CP
0,7558	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Age
0,7796	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak
0,7779	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang
0,7865	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach
0,7567	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg
0,7546	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, FBS
0,7580	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Chol
0,7602	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Trestbps
0,7690	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, CP
0,7756	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Age
<b>0,7857</b>	<b>Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal</b>	<b>Oldpeak, Exang</b>
0,7970	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach
0,7648	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg
0,7624	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak, FBS
0,7651	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Chol
0,7669	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Trestbps
0,7767	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak, CP
0,7830	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Age
0,7820	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Thalach
0,7531	Age, CP, Trestbps, Chol, FBS, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg
0,7512	Age, CP, Trestbps, Chol, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, FBS
0,7532	Age, CP, Trestbps, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Chol
0,7580	Age, CP, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Trestbps
0,7584	Age, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, CP
0,7745	CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Age
0,7655	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg
0,7630	Age, CP, Trestbps, Chol, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, FBS
0,7660	Age, CP, Trestbps, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Chol
0,7705	Age, CP, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Trestbps
0,7721	Age, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, CP
0,7781	CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Age
0,7287	Age, CP, Trestbps, Chol, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, FBS
0,7304	Age, CP, Trestbps, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Chol
0,7338	Age, CP, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Trestbps
0,7418	Age, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, CP
0,7469	CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Age
0,7288	Age, CP, Trestbps, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	FBS, Chol
0,7319	Age, CP, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	FBS, Trestbps
0,7403	Age, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	FBS, CP
0,7450	CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	FBS, Age
0,7338	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, Trestbps
0,7415	Age, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, CP
0,7463	CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, Age
0,7472	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Trestbps, CP
0,7486	CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Trestbps, Age
0,7610	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	CP, Age

Pada tabel 5. merupakan hasil dari proses metode Principal Component Analysis (PCA) dengan mereduksi 2 atribut dan kombinasi 10 atribut. Hasil yang didapatkan kombinasi atribut yang memiliki nilai Eigenvalue / perangkangan tertinggi terdapat pada kombinasi atribut Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal dengan mereduksi atribut Oldpeak dan Exang. Setelah didapatkan kombinasi atribut dengan nilai perangkangan tertinggi, selanjutnya melakukan klasifikasi kembali pada dataset dengan menggunakan kombinasi 10 atribut tersebut, sehingga didpatka hasil klasifikasi di bawah ini :

**Tabel 6.** Tingkat Akurasi Kombinasi 10 Atribut

	True CAD	True Normal	Class Precision
<b>Pred. CAD</b>	147	31	82,58%
<b>Pred. Normal</b>	18	107	85,60%
<b>Class Recall</b>	89,09%	77,54%	<b>83,83%</b>

Hasil pada tabel 6. terdapat tingkat persentase Precision, Recall dan Accuracy. Dari proses klasifikasi kombinasi 10 atribut terdapat akurasi yang berbeda dengan tingkat akurasi sebesar **83,83%**. Selanjutnya setelah



dilakukan reduksi atribut serta mengukur tingkat akurasi pada dataset dengan reduksi 2 atribut dan kombinasi 10 atribut, melakukan kembali proses reduksi dengan 3 atribut dan mencari kombinasi 9 atribut berdasarkan nilai Eigenvalue untuk melakukan perangkingan kombinasi atribut. Adapun hasil metode Principal Component Analysis (PCA) kombinasi 9 atribut dapat dilihat di bawah ini :

**Tabel 7.** Hasil Metode PCA Kombinasi 9 Atribut

Eigenvalue	Kombinasi	Reduksi
0,7556	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak	Thal, Ca, Slope
0,7586	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope	Thal, Ca, Oldpeak
0,7463	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope	Thal, Ca, Exang
0,7611	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope	Thal, Ca, Thalach
0,7217	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Restecg
0,7192	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Fbs
0,7211	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Chol
0,7252	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Trestbps
0,7318	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Cp
0,7385	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope	Thal, Ca, Age
0,7641	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Ca	Thal, Slope, Oldpeak
0,7616	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Ca	Thal, Slope, Exang
0,7748	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Ca	Thal, Slope, Thalach
0,7405	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Ca	Thal, Slope, Restecg
0,7382	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope, FBS
0,7414	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope, Chol
0,7445	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope, Trestbps
0,7524	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope, Cp
0,7637	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca	Thal, Slope, Age
0,7682	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca	Thal, Oldpeak, Exang
0,7842	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Slope, Ca	Thal, Oldpeak, Thalach
0,7475	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Slope, Ca	Thal, Oldpeak, Restecg
0,7450	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak, FBS
0,7472	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak, Chol
0,7498	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak, Trestbps
0,7588	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak, Cp
0,7693	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Ca	Thal, Oldpeak, Age
0,7664	Age, CP, Trestbps, Chol, FBS, Restecg, Oldpeak, Slope, Ca	Thal, Exang, Thalach
0,7337	Age, CP, Trestbps, Chol, FBS, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Restecg
0,7317	Age, CP, Trestbps, Chol, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Fbs
0,7332	Age, CP, Trestbps, FBS, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Chol
0,7393	Age, CP, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Trestbps
0,7383	Age, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Cp
0,7588	CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca	Thal, Exang, Age
0,7510	Age, CP, Trestbps, Chol, FBS, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Restecg
0,7483	Age, CP, Trestbps, Chol, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Fbs
0,7507	Age, CP, Trestbps, FBS, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Chol
0,7570	Age, CP, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Trestbps
0,7561	Age, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Cp
0,7662	CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca	Thal, Thalach, Age
0,7083	Age, CP, Trestbps, Chol, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg, Fbs
0,7097	Age, CP, Trestbps, FBS, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg, Chol
0,7139	Age, CP, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg, Trestbps
0,7216	Age, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg, Cp
0,7296	CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Restecg, Age
0,7080	Age, CP, Trestbps, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Fbs, Chol
0,7118	Age, CP, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Fbs, Trestbps
0,7201	Age, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Fbs, Cp
0,7275	CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Fbs, Age
0,7135	Age, CP, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Chol, Trestbps
0,7208	Age, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Chol, Cp
0,7284	CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Chol, Age
0,7277	Age, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Trestbps, Cp
0,7312	CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Trestbps, Age
0,7440	Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca	Thal, Cp, Age
0,7740	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Thal	Ca, Slope, Oldpeak
0,7760	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Thal	Ca, Slope, Exang
0,7829	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Thal	Ca, Slope, Thalach
0,7487	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Restecg
0,7458	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Fbs
0,7502	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Chol





Eigenvalue	Kombinasi	Reduksi
0,7521	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Trestbps
0,7628	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Cp
0,7664	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Thal	Ca, Slope, Age
0,7811	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Thal	Ca, Oldpeak, Exang
0,7916	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Slope, Thal	Ca, Oldpeak, Thalach
0,7545	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Restecg
0,7515	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Fbs
0,7548	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Chol
0,7563	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Trestbps
0,7678	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Cp
0,7712	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Slope, Thal	Ca, Oldpeak, Age
0,7753	Age, CP, Trestbps, Chol, FBS, Restecg, Oldpeak, Slope, Thal	Ca, Exang, Thalach
0,7428	Age, CP, Trestbps, Chol, FBS, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Restecg
0,7402	Age, CP, Trestbps, Chol, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Fbs
0,7427	Age, CP, Trestbps, FBS, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Chol
0,7482	Age, CP, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Trestbps
0,7477	Age, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Cp
0,7636	CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Thal	Ca, Exang, Age
0,7550	Age, CP, Trestbps, Chol, FBS, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Restecg
0,7517	Age, CP, Trestbps, Chol, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Fbs
0,7555	Age, CP, Trestbps, FBS, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Chol
0,7602	Age, CP, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Trestbps
0,7617	Age, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Cp
0,7656	CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Thal	Ca, Thalach, Age
0,7131	Age, CP, Trestbps, Chol, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg, Fbs
0,7152	Age, CP, Trestbps, FBS, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg, Chol
0,7187	Age, CP, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg, Trestbps
0,7276	Age, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg, Cp
0,7309	CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Restecg, Age
0,7132	Age, CP, Trestbps, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Fbs, Chol
0,7165	Age, CP, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Fbs, Trestbps
0,7255	Age, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Fbs, Cp
0,7288	CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Fbs, Age
0,7187	Age, CP, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Chol, Trestbps
0,7272	Age, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Chol, Cp
0,7303	CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Chol, Age
0,7334	Age, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Trestbps, Cp
0,7325	CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Trestbps, Age
0,7461	Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Thal	Ca, Cp, Age
0,7794	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Ca, Thal	Slope, Oldpeak, Exang
0,7880	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Ca, Thal	Slope, Oldpeak, Thalach
0,7591	Age, CP, Trestbps, Chol, FBS, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Restecg
0,7565	Age, CP, Trestbps, Chol, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Fbs
0,7610	Age, CP, Trestbps, FBS, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Chol
0,7613	Age, CP, Chol, FBS, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Trestbps
0,7727	Age, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Cp
0,7802	CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Ca, Thal	Slope, Oldpeak, Age
0,7827	Age, CP, Trestbps, Chol, FBS, Restecg, Oldpeak, Ca, Thal	Slope, Exang, Thalach
0,7573	Age, CP, Trestbps, Chol, FBS, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Restecg
0,7551	Age, CP, Trestbps, Chol, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Fbs
0,7593	Age, CP, Trestbps, FBS, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Chol
0,7635	Age, CP, Chol, FBS, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Trestbps
0,7631	Age, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Cp
0,7868	CP, Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Ca, Thal	Slope, Exang, Age
0,7678	Age, CP, Trestbps, Chol, FBS, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Restecg
0,7647	Age, CP, Trestbps, Chol, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Fbs
0,7706	Age, CP, Trestbps, FBS, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Chol
0,7735	Age, CP, Chol, FBS, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Trestbps
0,7751	Age, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Cp
0,7835	CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Ca, Thal	Slope, Thalach, Age
0,7305	Age, CP, Trestbps, Chol, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg, Fbs
0,7339	Age, CP, Trestbps, FBS, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg, Chol
0,7363	Age, CP, Chol, FBS, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg, Trestbps
0,7474	Age, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg, Cp
0,7532	CP, Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Restecg, Age
0,7321	Age, CP, Trestbps, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Fbs, Chol
0,7340	Age, CP, Chol, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Fbs, Trestbps



Eigenvalue	Kombinasi	Reduksi
0,7458	Age, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Fbs, Cp
0,7510	CP, Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Fbs, Age
0,7376	Age, CP, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Chol, Trestbps
0,7487	Age, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Chol, Cp
0,7536	CP, Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Chol, Age
0,7547	Age, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Trestbps, Cp
0,7545	CP, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Trestbps, Age
0,7742	Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Ca, Thal	Slope, Cp, Age
0,7933	Age, CP, Trestbps, Chol, FBS, Restecg, Slope, Ca, Thal	Oldpeak, Exang, Thalach
0,7669	Age, CP, Trestbps, Chol, FBS, Thalach, Slope, Ca, Thal	Oldpeak, Exang, Restecg
0,7645	Age, CP, Trestbps, Chol, Restecg, Thalach, Slope, Ca, Thal	Oldpeak, Exang, Fbs
0,7675	Age, CP, Trestbps, FBS, Restecg, Thalach, Slope, Ca, Thal	Oldpeak, Exang, Chol
0,7712	Age, CP, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal	Oldpeak, Exang, Trestbps
0,7713	Age, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal	Oldpeak, Exang, Cp
<b>0,7967</b>	<b>CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal</b>	<b>Oldpeak, Exang, Age</b>
0,7808	Age, CP, Trestbps, Chol, FBS, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Restecg
0,7772	Age, CP, Trestbps, Chol, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Fbs
0,7825	Age, CP, Trestbps, FBS, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Chol
0,7846	Age, CP, Chol, FBS, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Trestbps
0,7862	Age, Trestbps, Chol, FBS, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Cp
0,7954	CP, Trestbps, Chol, FBS, Restecg, Exang, Slope, Ca, Thal	Oldpeak, Thalach, Age
0,7398	Age, CP, Trestbps, Chol, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg, Fbs
0,7423	Age, CP, Trestbps, FBS, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg, Chol
0,7442	Age, CP, Chol, FBS, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg, Trestbps
0,7568	Age, Trestbps, Chol, FBS, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg, Cp
0,7619	CP, Trestbps, Chol, FBS, Thalach, Exang, Slope, Ca, Thal	Oldpeak, Restecg, Age
0,7631	Age, CP, Trestbps, Chol, FBS, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Restecg
0,7606	Age, CP, Trestbps, Chol, Restecg, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Fbs
0,7638	Age, CP, Trestbps, FBS, Restecg, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Chol
0,7716	Age, CP, Chol, FBS, Restecg, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Trestbps
0,7637	Age, Trestbps, Chol, FBS, Restecg, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Cp
0,7819	CP, Trestbps, Chol, FBS, Restecg, Oldpeak, Slope, Ca, Thal	Exang, Thalach, Age
0,7270	Age, CP, Trestbps, Chol, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg, Fbs
0,7288	Age, CP, Trestbps, FBS, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg, Chol
0,7341	Age, CP, Chol, FBS, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg, Trestbps
0,7355	Age, Trestbps, Chol, FBS, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg, Cp
0,7523	CP, Trestbps, Chol, FBS, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Restecg, Age
0,7272	Age, CP, Trestbps, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Fbs, Chol
0,7319	Age, CP, Chol, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Fbs, Trestbps
0,7341	Age, Trestbps, Chol, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Fbs, Cp
0,7501	CP, Trestbps, Chol, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Fbs, Age
0,7341	Age, CP, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Chol, Trestbps
0,7353	Age, Trestbps, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Chol, Cp
0,7514	CP, Trestbps, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Chol, Age
0,7432	Age, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Trestbps, Cp
0,7551	CP, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Trestbps, Age
0,7612	Trestbps, Chol, FBS, Restecg, Thalach, Oldpeak, Slope, Ca, Thal	Exang, Cp, Age
0,7408	Age, CP, Trestbps, Chol, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg, Fbs
0,7437	Age, CP, Trestbps, FBS, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg, Chol
0,7487	Age, CP, Chol, FBS, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg, Trestbps
0,7518	Age, Trestbps, Chol, FBS, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg, Cp
0,7569	CP, Trestbps, Chol, FBS, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Restecg, Age
0,7416	Age, CP, Trestbps, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Fbs, Chol
0,7458	Age, CP, Chol, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Fbs, Trestbps
0,7499	Age, Trestbps, Chol, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Fbs, Cp
0,7542	CP, Trestbps, Chol, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Fbs, Age
0,7492	Age, CP, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Chol, Trestbps
0,7520	Age, Trestbps, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Chol, Cp
0,7565	CP, Trestbps, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Chol, Age
0,7607	Age, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Trestbps, Cp
0,7596	CP, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Trestbps, Age
0,7682	Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	Thalach, Cp, Age
0,7013	Age, CP, Trestbps, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Fbs, Chol
0,7047	Age, CP, Chol, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Fbs, Trestbps
0,7148	Age, Trestbps, Chol, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Fbs, Cp
0,7191	CP, Trestbps, Chol, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Fbs, Age
0,7065	Age, CP, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Chol, Trestbps



Eigenvalue	Kombinasi	Reduksi
0,7159	Age, Trestbps, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Chol, Cp
0,7203	CP, Trestbps, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Chol, Age
0,7221	Age, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Trestbps, Cp
0,7227	CP, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Trestbps, Age
0,7373	Trestbps, Chol, FBS, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Restecg, Cp, Age
0,7047	Age, CP, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Chol, Trestbps
0,7147	Age, Trestbps, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Chol, Cp
0,7186	CP, Trestbps, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Chol, Age
0,7203	Age, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Trestbps, Cp
0,7207	CP, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Trestbps, Age
0,7356	Trestbps, Chol, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Fbs, Cp, Age
0,7218	Age, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, Trestbps, Cp
0,7221	CP, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, Trestbps, Age
0,7362	Trestbps, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Chol, Cp, Age
0,7410	Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	Trestbps, Cp, Age

Tabel 7. merupakan hasil dari proses metode Principal Component Analysis (PCA) dengan mereduksi 3 atribut dan kombinasi 9 atribut. Dari hasil proses tersebut didapatkan kombinasi atribut yang memiliki nilai Eigenvalue atau perangkingan kombinasi atribut tertinggi terdapat pada kombinasi atribut CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal dengan mereduksi atribut Oldpeak, Exang dan Age. Setelah diketahui kombinasi atribut dengan nilai Eigenvalue tertinggi, kemudian melakukan klasifikasi kembali dengan menggunakan kombinasi 9 atribut tersebut. Sehingga didapatkan hasil klasifikasi seperti di bawah ini :

**Tabel 8.** Tingkat Akurasi Kombinasi 9 Atribut

	True CAD	True Normal	Class Precision
<b>Pred. CAD</b>	149	34	81,42%
<b>Pred. Normal</b>	16	104	86,67%
<b>Class Recall</b>	90,30%	75,36%	<b>83,50%</b>

Pada tabel 8. yang merupakan hasil tingkat akurasi memiliki persentasi Precision, Recall dan Accuracy. Dari hasil proses kombinasi 9 atribut tersebut, terdapat akurasi yang berbeda dari hasil proses klasifikasi sebelumnya dengan tingkat akurasi sebesar **83,50%**. Dari hasil tingkat akurasi reduksi 3 atribut dengan kombinasi 9 atribut terdapat perbedaan tingkat akurasi dengan hasil klasifikasi pada kombinasi 11 dan 10 atribut. Pada dataset dilakukan reduksi 3 atribut Oldpeak, Exang dan Age dengan kombinasi 9 atribut CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal tingkat akurasi yang didapatkan sebesar **83,50%**. Hal ini dikarenakan terdapat penurunan pada prediksi dengan hasil klasifikasi Normal, dimana pada hasil klasifikasi reduksi 2 atribut dengan kombinasi 10 atribut untuk atribut Oldpeak Exang dan Age digunakan menjadi pohon pengambilan keputusan pada penentuan proses hasil klasifikasi. Sedangkan pada reduksi 3 atribut dengan kombinasi 9 atribut, Oldpeak dan Exang mengalami reduksi atribut berdasarkan kombinasi nilai Eigenvalue pada metode PCA. Reduksi atribut yang menjadi pohon keputusan ini menyebabkan berkurangnya tingkat akurasi pada proses klasifikasi terkhususnya pada hasil klasifikasi Normal.

Setelah dilakukan proses reduksi atribut pada dataset dan mengetahui kombinasi – kombinasi atribut dengan nilai Eigenvalue / perangkingan tertinggi dengan menggunakan metode Principal Componen Analysis (PCA) serta melakukan klasifikasi dan mengukur tingkat akurasi pada proses klasifikasi, selanjutnya adalah menampilkan seluruh hasil proses pengujian. Untuk hasil pengujian dataset dapat dilihat pada tabel di bawah ini :

**Tabel 9.** Hasil Pengujian Dataset

No	Reduksi Atribut	Kombinasi Atribut	Tingkat Akurasi
1	-	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal	87,46%
2	Thalach	Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal	89,11%
3	Oldpeak, Exang	Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal	83,83%
4	Oldpeak, Exang, Age	CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal	83,50%

### 3.2 Pembahasan

Pada proses pengujian dataset dengan melakukan reduksi atribut dan melakukan klasifikasi dengan kombinasi atribut yang berbeda – beda didapatkan tingkat akurasi yang berbeda juga. Tingkat akurasi pada dataset sebelum dilakukan reduksi atribut didapat akurasi sebesar 87,46%, setelah dilakukan reduksi 1 atribut (Thalach) dengan kombinasi atribut Age, CP, Trestbps, Chol, FBS, Restecg, Exang, Oldpeak, Slope, Ca, Thal didapatkan tingkat akurasi sebesar 89,11%, pada reduksi 2 atribut (Oldpeak dan Exang) dengan kombinasi atribut Age, CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope, Ca, Thal tingkat akurasi yang didapatkan sebesar 83,83% dan reduksi 3 atribut (Oldpeak, Exang dan Age) dengan kombinasi atribut CP, Trestbps, Chol, FBS, Restecg, Thalach, Slope,



Ca, Thal didapatkan akurasi sebesar 83,50%. Setelah dilakukan reduksi 1 atribut dengan kombinasi 11 atribut setiap dilakukan reduksi atribut mengalami penurunan pada tingkat akurasi hal ini dikarenakan ketika dilakukan proses klasifikasi sebelum terdapat reduksi atribut seluruh atribut pada dataset menjadi rule ataupun pohon keputusan, begitu juga ketika proses klasifikasi pada reduksi 1 atribut dengan kombinasi 11 atribut yang kemudian dilakukan proses klasifikasi. Seluruh kombinasi atribut tersebut menjadi pohon keputusan juga, sehingga ketika dilakukan reduksi atribut akan mempengaruhi tingkat akurasi pada proses klasifikasi lainnya.

Hal tersebut yang mempengaruhi pada proses pengambilan keputusan hasil klasifikasi, dimana terdapat penurunan pada kebenaran dalam menemukan sebuah hasil proses klasifikasi atau Class Recall Prediction dan ketepatan informasi yang diberikan atau Class Precision. Dari proses reduksi dan klasifikasi didapatkan bahwa pada dataset untuk atribut optimal terdapat saat sebelum dilakukannya reduksi atribut, dikarenakan hasil tingkat akurasi yang tertinggi terdapat pada saat tidak dilakukannya reduksi pada atribut dataset.

#### 4. KESIMPULAN

Berdasarkan dari penjelasan permasalahan dan uraian setiap BAB serta hasil dari penelitian, maka dapat ditarik kesimpulan (a) Metode Principal Componen Analysis (PCA) dapat digunakan untuk menentukan perankingan kombinasi atribut yang memiliki nilai kombinasi tertinggi berdasarkan nilai Eigenvalue. (b) Proses klasifikasi menggunakan algoritma C5.0 pada dataset menghasilkan pohon keputusan dan tingkat akurasi yang berbeda – beda pada setiap reduksi atribut dengan tingkat akurasi yang paling rendah pada reduksi 3 atribut dengan memiliki akurasi sebesar 83,50% dan paling tinggi pada sebelum dilakukan reduksi atribut dengan akurasi sebesar 89,11%, hal ini dikarenakan yang menjadi pohon keputusan mengalami reduksi. (c) Pada dataset atribut optimal terdapat ketika tidak dilakukannya reduksi atribut dengan tingkat akurasi tertinggi pada sebelum dilakukannya reduksi atribut pada dataset.

#### REFERENCES

- [1] F. Babič, J. Olejár, Z. Vantová and J. Paralič, "Predictive and Descriptive Analysis for Heart Disease Diagnosis," *Federated Conference on Computer Science and Information Systems*, vol. 11, pp. 155-163, 2017.
- [2] D. B. Umadevi and M. Snehapriya, "A Survey on Prediction of Heart Disease Using Data Mining Techniques," *International Journal of Science and Research (IJSR)*, vol. 6, no. 4, pp. 2228-2232, 2017.
- [3] J. H. B. B. H. M. Roohallah Alizadehsani, A. Ghandeharioun, Reihane Boghrati and Z. A. Sani, "Diagnosis Of Coronary Arteries Stenosis Using Data Mining," *Journal of Medical Signals & Sensors*, vol. 2, no. 3, pp. 153-160, 2012.
- [4] D. Chaki, A. Das and M. I. Zaber, "A Comparison of Three Discrete Methods for Classification of Heart Disease Data," *Bangladesh Journal Of Scientific And Industrial Research*, vol. 50, no. 4, pp. 293-296, 2015.
- [5] B. Kaur and W. Singh, "Review on Heart Disease Prediction System using Data Mining Techniques," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 2, no. 10, pp. 3003-3008, 2014.
- [6] R. Alizadehsani, J. Habibi, M. J. Hosseini, H. Mashayekhi, R. Bogharti, A. Ghandeharioun, B. Bahadorian and Z. A. Sani, "A Data Mining Approach For Diagnosis Of Coronary Artery Disease," *Computer Methods and Programs In Biomedicine*, pp. 1-10, 2013.
- [7] R. El-Bialy, M. A. Salamay, O. H. Karam and M. Khalifa, "Feature Analysis of Coronary Artery Heart Disease Data Sets," *International Conference on Communication, Management and Information Technology*, pp. 459-468, 2015.
- [8] D. K. B. A. Janabi and R. Kadhim, "Data Reduction Techniques: A Comparative Study for Attribute Selection Methods," *International Journal of Advanced Computer Science and Technology*, vol. 8, no. 1, pp. 1-13, 2018.
- [9] W. Ding, J. Wang and Z. Guan, "A Novel Minimum Attribute Reduction Algorithm Based on Hierarchical Elitist Role Model Combining Competitive and Cooperative Co-evolution," *Chinese Journal of Electronics*, vol. 22, no. 4, pp. 677-682, 2013.
- [10] I. T. Jolliffe and J. Cadima, "Principal Component Analysis: A Review And Recent Developments," *The Royal Society Publishing*, pp. 1-16, 2016.
- [11] A. R. Syakhala, D. Puspitaningrum and E. P. Purwandari, "Perbandingan Metode Principal Component Analysis (Pca) Dengan Metode Hidden Markov Model (Hmm) Dalam Pengenalan Identitas Seseorang Melalui Wajah," *Jurnal Rekursif*, vol. 3, no. 2, pp. 68-81, 2015.
- [12] M. Abdar, S. R. N. Kalhori, T. Sutikno, I. M. I. Subroto and G. Arji, "Comparing Performance of Data Mining Algorithms In Prediction Heart Diseases," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 6, pp. 1569-1576, 2015.
- [13] R. Pandya and J. Pandya, "C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning," *International Journal of Computer Applications*, vol. 117, no. 16, pp. 18-21, 2015.
- [14] R. Revathy and R. Lawrance, "Comparative Analysis of C4.5 and C5.0 Algorithms on Crop Pest Data," *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)*, vol. 5, no. 1, pp. 50-58, 2017.
- [15] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques*, Third ed., USA: Morgan Kaufmann, 2012.
- [16] K. V. K and S. B, "Dimensionality Reduction Using Principal Component Analysis For Network Intrusion Detection," *Perspectives in Science*, vol. 8, pp. 510-512, 2016.
- [17] S.-H. Wang, T.-M. Zhan, Y. Chen, Y. Zhang, M. Yang, H.-M. Lu, H.-N. Wang, B. Liu and P. Phillips, "Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression," *IEEE*, vol. 4, pp. 7567-7576, 2016.
- [18] D. Nandi, A. S. Ashour, S. Samanta, S. Chakraborty, M. A. Salem and N. Dey, "Principal Component Analysis In Medical Image Processing: A Study," *International Journal of Image Mining*, vol. 1, no. 1, pp. 65-86, 2015.



- [19] M. Abdar, M. Zomorodi-Moghadam, R. Das and I.-H. Ting, "Performance Analysis Of Classification Algorithms On Early Detection Of Liver Disease," *Expert Systems With Applications* 125, pp. 442-443, 2019.
- [20] B. R. Patel and K. K. Rana, "A Survey on Decision Tree Algorithm For Classification," *International Journal of Engineering Development and Research*, vol. 2, no. 1, pp. 1-5, 2014.
- [21] I. H. Witten, E. Frank and M. A. Hall, *Data Mining Practical Machine Learning Tools and Techniques*, Third ed., USA: Morgan Kaufmann, 2011.
- [22] I. H. Witten and E. Frank, *Data Mining Practical Machine Learning Tools and Techniques*, Second ed., USA: Morgan Kaufmann, 2005.
- [23] C. M. Bishop, *Pattern Recognition and Machine Learning*, First ed., Singapore: Business Media, 2006.