

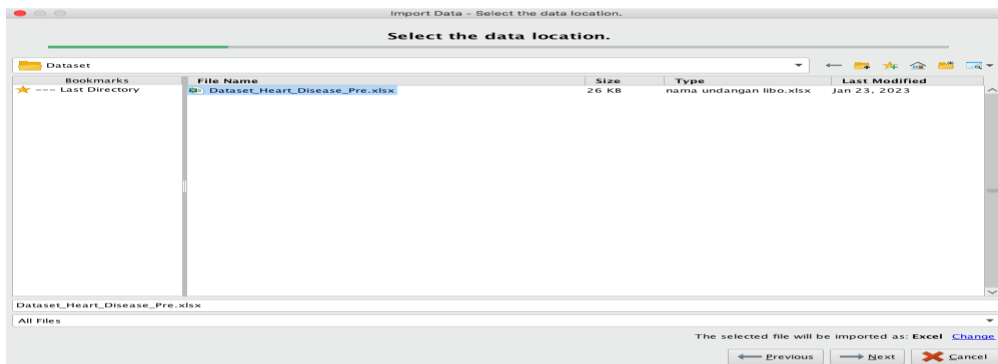
BAB IV HASIL DAN PEMBAHASAN

1.1 Hasil Penelitian

Pada penelitian ini digunakan seleksi fitur *PSO* dan *Algoritma Naive Bayes*. kasus yang diambil adalah data penyakit jantung yang berjumlah 303 data. Data sample ini yang akan dihitung probabilitasnya, kemudian akan dicari tingkat akurasi tertinggi dari data tersebut dan akan diimplementasikan pada *tools RapidMiner*.

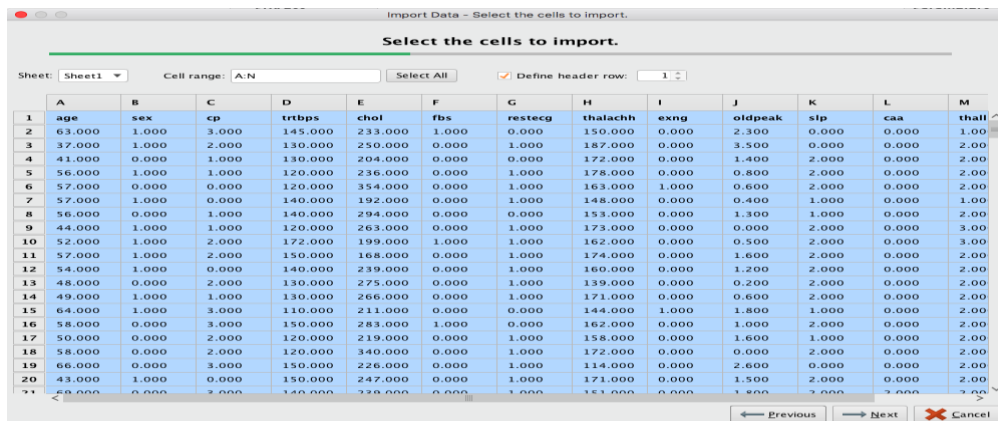
1.2 Dataset

Dataset berupa excel yang telah siap dan sudah melalui tahap *preprocessing* akan di *import* ke dalam *tools rapidminer*, ambil dataset dari penyimpanan komputer lalu pilih dan *import*.



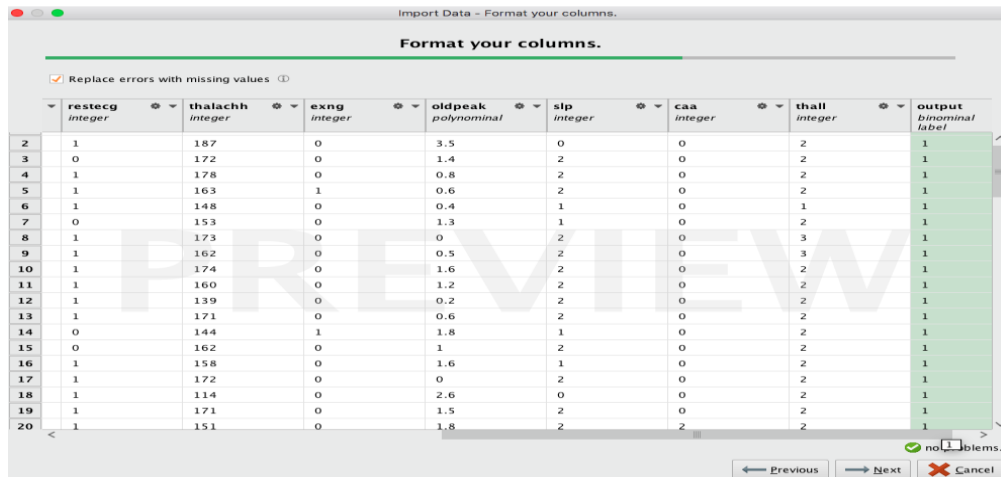
Gambar 4. 1 *Import Dataset*

Kemudian setelah di pilih *next* maka muncul tampilan seperti gambar berikut ini. Pastikan nama *sheet* dan *cell range dataset* sudah sesuai dengan data yang dimiliki.



Gambar 4. 2 Kolom Dataset

Kemudian atur *type atribut* sesuai dengan *value* atau nilai yang ada didalamnya serta atur *role atribut* seperti *id* dan label pada dataset seperti gambar berikut ini.

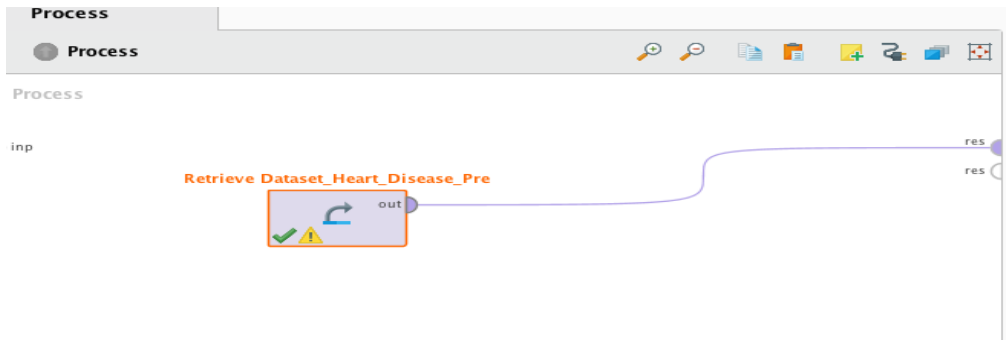


Gambar 4. 3 Format Type Dataset

Tabel 4. 1 Type Data Atribut

No	Atribut	Type Data
1	age	Integer
2	sex	Binominal
3	cp	Polynomial
4	trtbps	Integer
5	chol	Integer
6	fbs	Binominal
7	restecg	Binominal
8	thalachh	Integer
9	exng	Binominal
10	oldpeak	Polynomial
11	slp	Polynomial
12	caa	Polynomial
13	thall	Polynomial
14	output	Binominal

Untuk dapat melihat dataset pada *tools rapidminer* bisa dilakukasn seperti gambar dibawah ini.



Gambar 4. 4 Tampilan Dataset Di *Rapidminer*

Type atribut juga dapat dilihat dari statistics pada *tools rapidminer*, selain *type* data kita juga dapat melihat *missing* data dan nilai min dan max pada setiap atribut

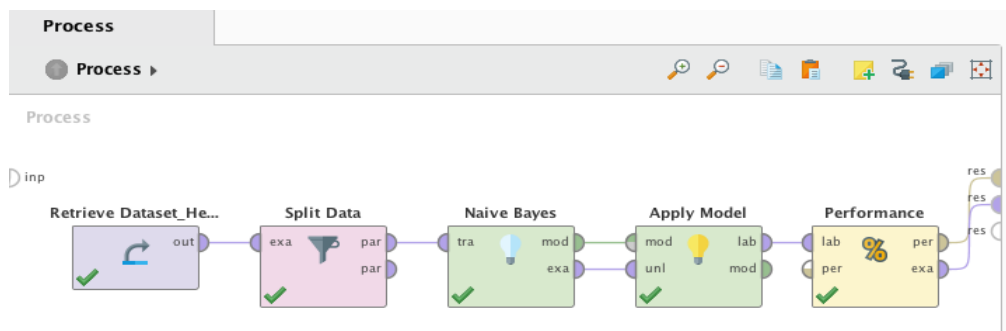
The screenshot shows the Results view in RapidMiner. The table displays the following data:

Name	Type	Missing	Statistics	Filter (14 / 14 attributes):
Label output	Binominal	0	Negative 1 Positive 0	Values 1 (165), 0 (138)
age	Integer	0	Min 29 Max 77	Average 54.366
sex	Binominal	0	Negative 1 Positive 0	Values 1 (207), 0 (96)
cp	Nominal	0	Least 3 (23) Most 0 (143)	Values 0 (143), 2 (87), ...[2]
trtbps	Integer	0	Min 94 Max 200	Average 131.624
chol	Integer	0	Min 126 Max 564	Average 246.264
fbs	Binominal	0	Negative 1 Positive 0	Values 0 (258), 1 (45)

Gambar 4. 5 Statistik Dataset

1.3 Algoritma *Naive Bayes* Dengan *Split Validation*

Penerapan *algoritma naive bayes* pada *rapidminer* dengan menggunakan *split validation* dengan nilai akurasi, *precision*, *confusion matrix* atau nilai *recall* dan nilai *AUC* dengan pembagian data *training* dan data *testing* sebesar 80:20, alur pada *rapidminer* dapat dilihat pada gambar berikut.



Gambar 4. 6 Proses *Naive Bayes* Dengan *Split Validation*

Hasil Akurasi *Naïve Bayes* dan *Split Validation*, pada pembagian data sampling 80% dan data testing 20% maka didapatkan akurasi sebesar 87,60%

Table View Plot View

accuracy: 87.60%

	true 1	true 0	class precision
pred. 1	119	17	87.50%
pred. 0	13	93	87.74%
class recall	90.15%	84.55%	

Gambar 4. 7 Nilai Akurasi Dengan *Split Validation*

Menghitung akurasi :

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

$$\text{Akurasi} = \frac{(119+93)}{(119+93+17+13)} = \frac{212}{242} = 87,60\%$$

Berikut adalah hasil *precision* yaitu mendapatkan 87,74% seperti gambar dibawah ini.

Table View Plot View

precision: 87.74% (positive class: 0)

	true 1	true 0	class precision
pred. 1	119	17	87.50%
pred. 0	13	93	87.74%
class recall	90.15%	84.55%	

Gambar 4. 8 Nilai *Precision* Dengan *Split Validation*

Menghitung *Precision* :

$$\text{Rumus Precision} = \frac{Tp}{Tp+Fp}$$

$$P(1) = \frac{119}{119+17} = \frac{119}{136} = 87,50\% \quad P(0) = \frac{93}{93+13} = \frac{93}{106} = 87,74\%$$

Berikut adalah hasil *confusion matrix* yaitu mendapatkan 84,55% seperti gambar dibawah ini.

Table View Plot View

recall: 84.55% (positive class: 0)

	true 1	true 0	class precision
pred. 1	119	17	87.50%
pred. 0	13	93	87.74%
class recall	90.15%	84.55%	

Gambar 4. 9 Nilai *Confusion Matrix* Dengan *Split Validation*

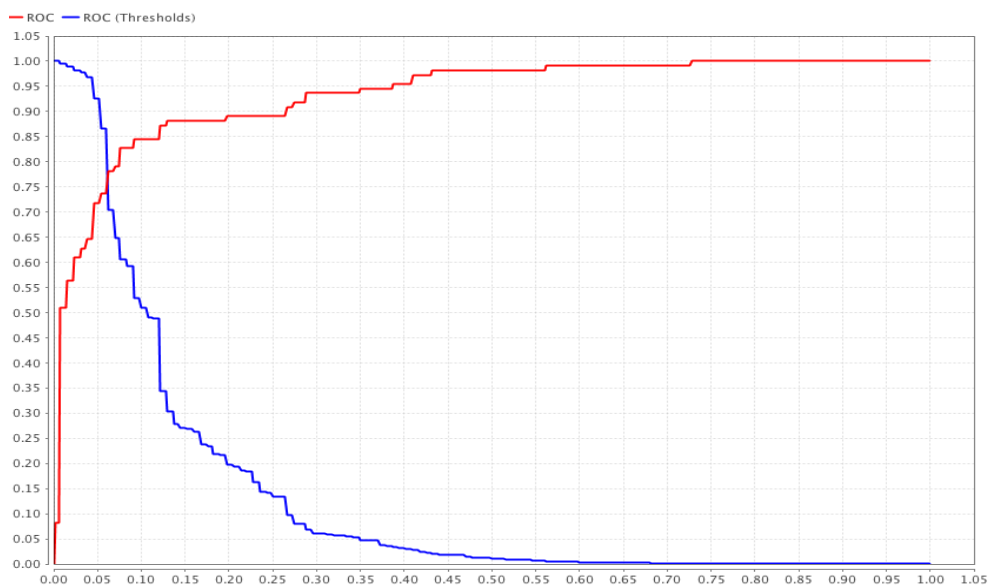
Menghitung *Recall* :

$$\text{Rumus Recall} = \frac{Tp}{Tp+Fn}$$

$$R(1) = \frac{119}{119+13} = \frac{119}{132} = 90,15 \% \quad R(0) = \frac{93}{93+17} = \frac{93}{110} = 84,55 \%$$

Berikut adalah hasil *kurva AUC* yaitu mendapatkan nilai 0,933 seperti gambar dibawah ini.

AUC: 0.933 (positive class: 0)



Gambar 4. 10 *Kurva AUC* Dengan *Split Validation*

Tabel 4. 2 Hasil dari *Naïve Bayes* dan *Split Validation* (80:20)

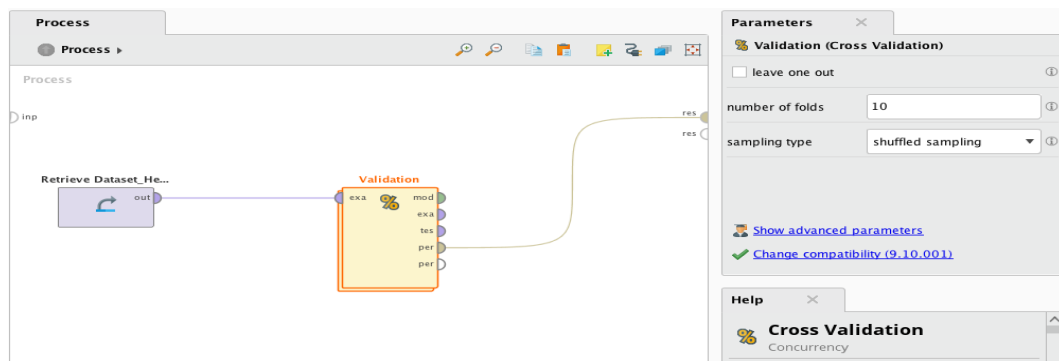
No	Naïve Bayes (80:20)	Hasil Nilai
1	<i>Accuracy</i>	87,60%
2	<i>Precision</i>	87,74%
3	<i>Recall / Confusion Matrix</i>	84,55%
4	<i>AUC</i>	0,933

1.4 Algoritma *Naïve Bayes* menggunakan *Cross Validation* (90:10)

Metode ini adalah untuk mengetahui dan membandingkan tingkat akurasi dalam klasifikasi dengan menggunakan *cross validation* yang di implementasikan dengan menggunakan *Naïve Bayes*. untuk perbandingan data training sebesar 90% dan testing 10%. penerapan metode ini dapat dilihat pada gambar dibawah ini.

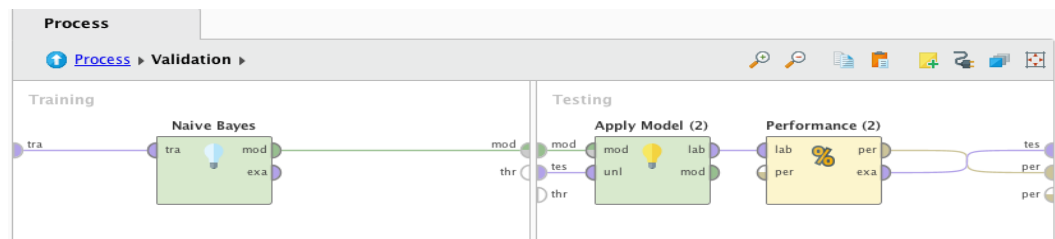
A. Tahapan Operator

Untuk penyelesaian dataset *heart disease* atau penyakit jantung *classification*, ditunjukkan oleh gambar berikut.



Gambar 4. 11. Proses *Naïve Bayes* dengan *Cross Validation*

Proses pertama yaitu dengan membaca file *dataset disease*, selanjutnya hubungkan dengan operator *Cross Validation*. Percobaan ini dilakukan terhadap dataset menggunakan *tools rapidminer* dengan *number of fold* 10, tahap selanjutnya dapat dilihat pada gambar dibawah ini.



Gambar 4. 12 Proses *Cross Validation*

Untuk dapat menampilkan tahap ini dengan cara mengklik 2x pada operator *Cross Validation*, kemudian akan tampil 2 jendela, yang pertama jendela training, pada jendela ini isikan dengan operator *Naïve Bayes*, selanjutnya pada jendela testing isikan dengan operator *Apply model* dan *Performance*.

B. Hasil Accuracy

Berikut adalah hasil *accuracy* yaitu mendapatkan 82.83% seperti gambar dibawah ini.

Table View Plot View

accuracy: 82.83% +/- 5.81% (micro average: 82.84%)

	true 1	true 0	class precision
pred. 1	141	28	83.43%
pred. 0	24	110	82.09%
class recall	85.45%	79.71%	

Gambar 4. 13 Hasil Akurasi *Naïve Bayes* dengan *Cross Validation*

C. Hasil Precision

Berikut adalah hasil *precision* yaitu mendapatkan 82,10% seperti gambar dibawah ini.

Table View Plot View

precision: 82.10% +/- 9.91% (micro average: 82.09%) (positive class: 0)

	true 1	true 0	class precision
pred. 1	141	28	83.43%
pred. 0	24	110	82.09%
class recall	85.45%	79.71%	

Gambar 4. 14 Hasil *Precision Naïve Bayes* dengan *Cross Validation*

D. Hasil Confusion Matrix

Berikut adalah hasil *confusion matrix* yaitu mendapatkan 79,17% seperti gambar dibawah ini.

Table View Plot View

recall: 79.17% +/- 11.85% (micro average: 79.71%) (positive class: 0)

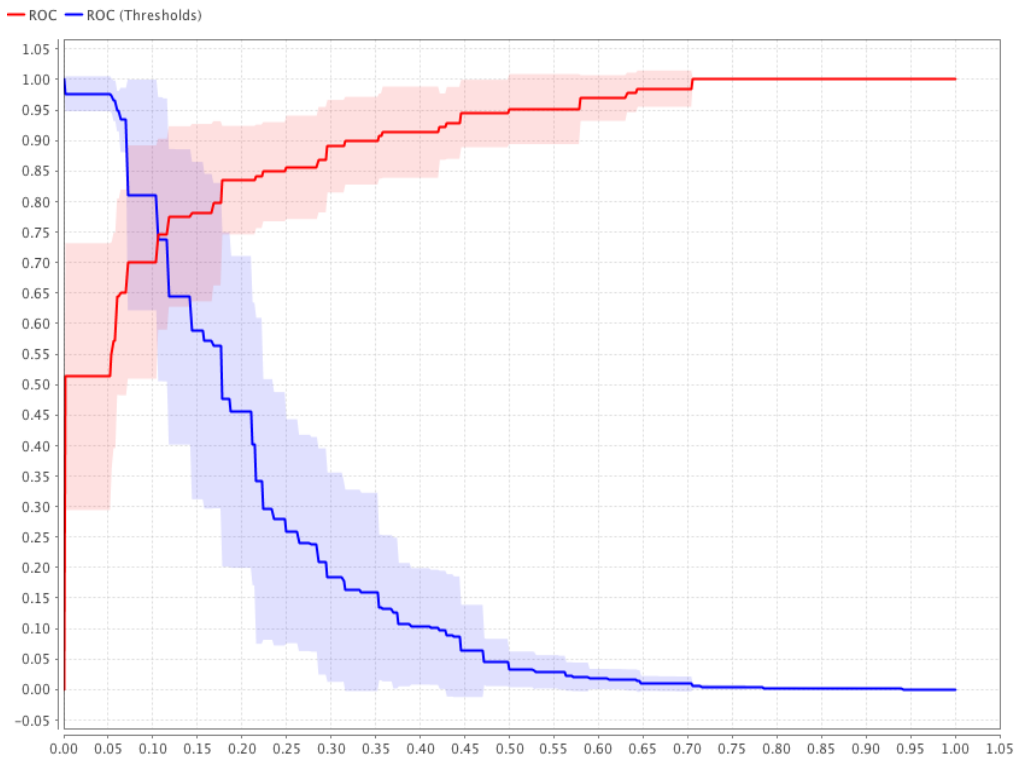
	true 1	true 0	class precision
pred. 1	141	28	83.43%
pred. 0	24	110	82.09%
class recall	85.45%	79.71%	

Gambar 4. 15 Hasil *Confusion Matrix Naïve Bayes* dengan *Cross Validation*

E. Hasil AUC

Berikut adalah hasil *kurva AUC* yaitu mendapatkan nilai 0,900 seperti gambar dibawah ini.

AUC: 0.900 +/- 0.053 (micro average: 0.900) (positive class: 0)



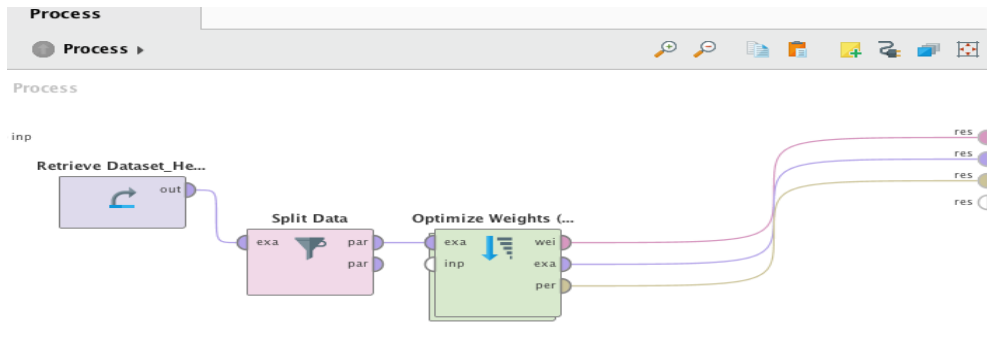
Gambar 4. 16 Gambar dari grafik AUC

Tabel 4. 3 Hasil dari *Naïve Bayes* dan *Cross Validation* (90:10)

No	Naïve Bayes (90:10)	Hasil Nilai
1	<i>Accuracy</i>	82.83%
2	<i>Precision</i>	82,10%
3	<i>Recall / Confusion Matrix</i>	79,17%
4	<i>AUC</i>	0,900

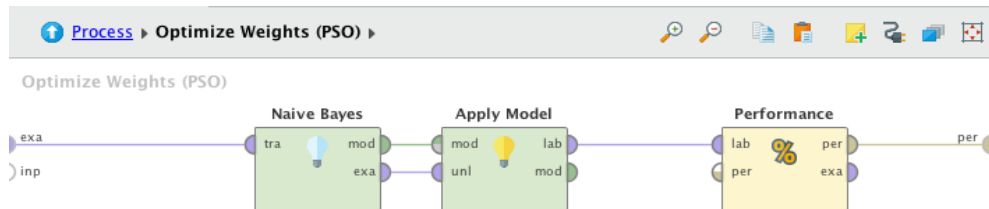
1.5 Pengujian *Patricle Swarm Optimization (PSO)* dan *Naïve Bayes* Dengan *Split Validation* (80:20)

Metode ini adalah untuk meningkatkan akurasi dalam klasifikasi dengan menggunakan Teknik seleksi fitur pembobotan atribut dengan menggunakan *PSO* dengan *Naïve Bayes* serta perbandingan *split data training* sebesar 80% dan *testing* 20%. penerapan metode ini dapat dilihat pada gambar dibawah ini.



Gambar 4. 17 Proses *PSO* dan *Naïve Bayes*

Setelah seperti gambar diatas kita masukkan *algoritma Naïve Bayes* kedalam *PSO* seperti gambar dibawah ini.



Gambar 4. 18 Sub Proses *PSO* dan *Naïve Bayes*

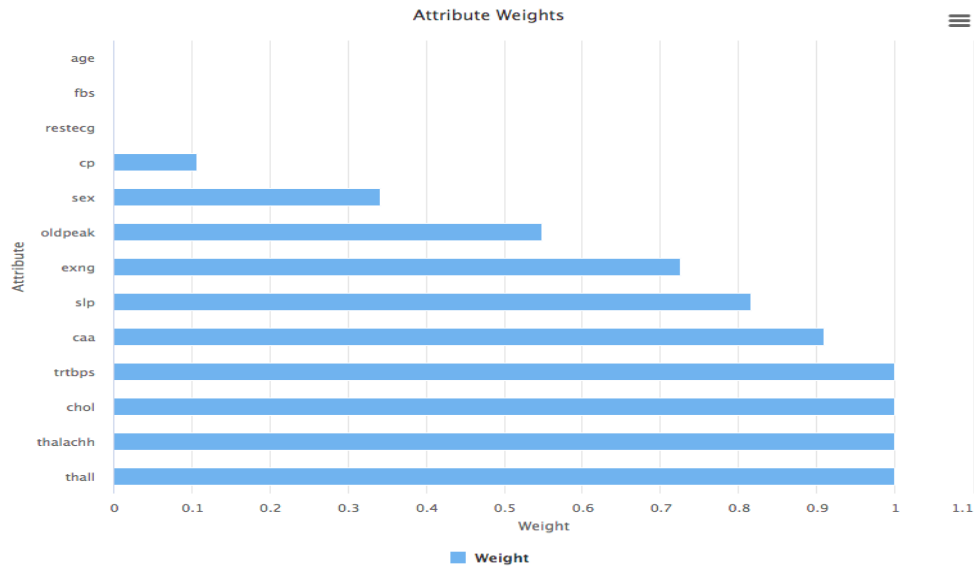
Atribut yang digunakan pada *algoritma Naïve Bayes* ini melibatkan semua atribut yang berjumlah 13, Adapun atribut dan bobot dapat dilihat dari tabel berikut.

Tabel 4. 4 Tabel Atribut Yang Digunakan Sesudah PSO

No	Atribut	Weight
1	age	0
2	sex	0.341
3	cp	0.106
4	trtbps	1
5	chol	1
6	fbs	0
7	restecg	0
8	thalachh	1
9	exng	0.726
10	oldpeak	0.548
11	slp	0.816

12	caa	0.909
13	thall	1

Dari keterangan tabel diatas bahwa terdapat atribut yang memiliki nilai 0 yaitu age, fbs, restecg yang artinya pada proses *PSO* atribut tersebut tidak berkontribusi dalam proses klasifikasi prediksi penyakit jantung dengan menggunakan algoritma *Naive Bayes*.

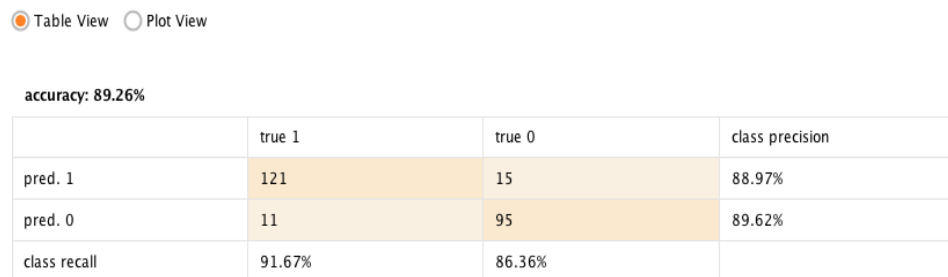


Gambar 4. 19 Gambar Pembobotan Atribut Dengan PSO

Dari pengaplikasian *PSO* dan *Naive Bayes* pada *rapidminer* maka dihasilkan peningkatan akurasi, dimana dapat dilihat perbandingan akurasi, *precision* dan *recall* serta *AUC* sesudah dan sebelum di tambahkan metode *PSO*.

A. Hasil Accuracy

Berikut adalah hasil *accuracy* yaitu mendapatkan 89,26% seperti gambar dibawah ini.



Gambar 4. 20 Nilai Akurasi *PSO* dan *Naive Bayes*

B. Hasil Precision

Berikut adalah hasil *precision* yaitu mendapatkan 89,62% seperti gambar dibawah ini.

Table View Plot View

precision: 89.62% (positive class: 0)

	true 1	true 0	class precision
pred. 1	121	15	88.97%
pred. 0	11	95	89.62%
class recall	91.67%	86.36%	

Gambar 4. 21 Nilai *Precision PSO* dan *Naïve Bayes*

C. Hasil *Confusion Matrix*

Berikut adalah hasil *confusion matrix* yaitu mendapatkan 86,36% seperti gambar dibawah ini.

Table View Plot View

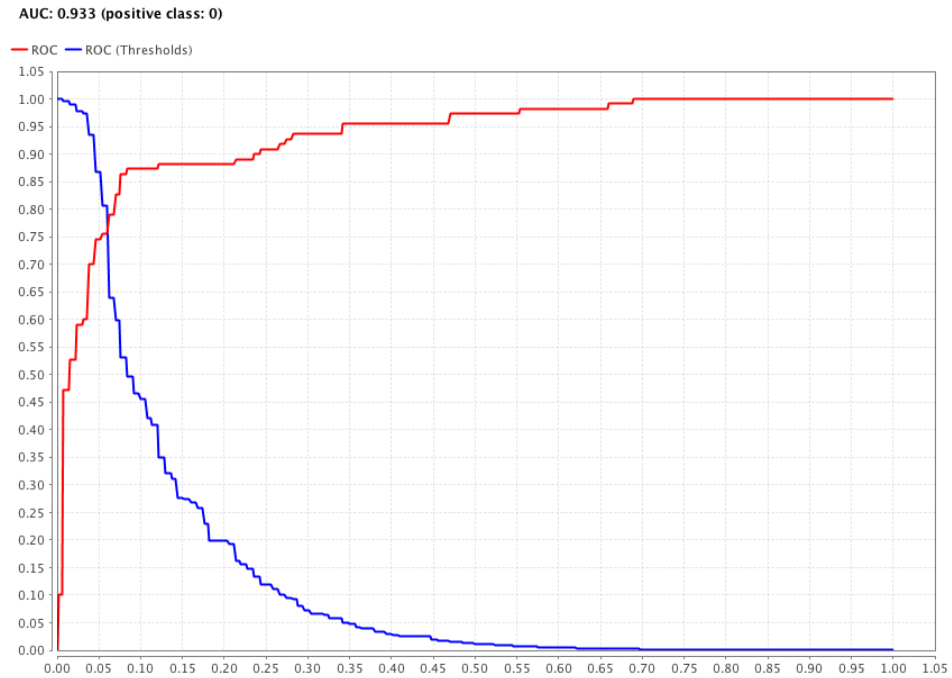
recall: 86.36% (positive class: 0)

	true 1	true 0	class precision
pred. 1	121	15	88.97%
pred. 0	11	95	89.62%
class recall	91.67%	86.36%	

Gambar 4. 22 Nilai *Confusion Matrix PSO* dan *Naïve Bayes*

D. Hasil *AUC*

Berikut adalah hasil *kurva AUC* yaitu mendapatkan nilai 0,933 seperti gambar dibawah ini.



Gambar 4. 23 Kurva AUC PSO dan Naïve Bayes

Tabel 4. 5 Hasil dari PSO dan Naïve Bayes (80:20)

No	PSO + Naïve Bayes (80:20)	Hasil Nilai
1	<i>Accuracy</i>	89,26%
2	<i>Precision</i>	89,62%
3	<i>Recall / Confusion Matrix</i>	86,36%
4	<i>AUC</i>	0,933

1.6 Pembahasan Hasil Penelitian

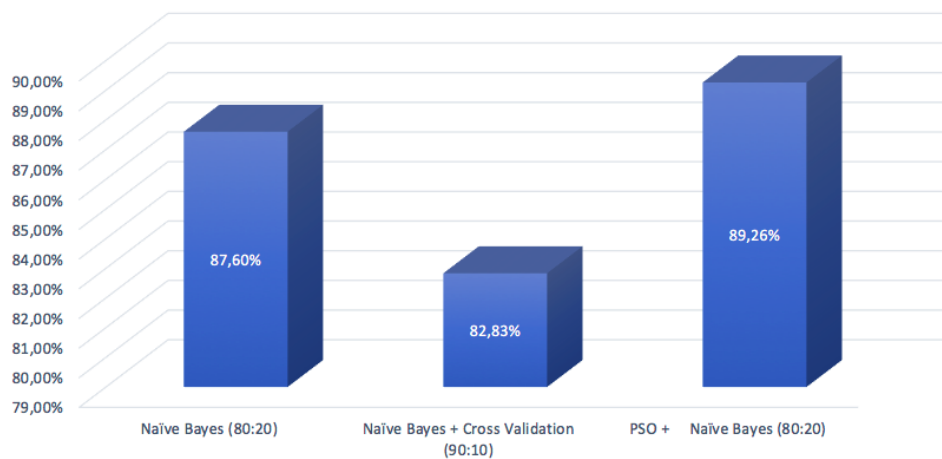
Pembahasan hasil penelitian yaitu membahas hasil penelitian yang telah dilakukan. Penelitian yang dilakukan yang pertama yaitu mengklasifikasi penyakit jantung dengan menggunakan *Algoritma Naïve Bayes* dan penelitian ke dua yaitu untuk menabah akurasi sehingga ditambahkan metode seleksi fitur *PSO* dan *Algoritma Naïve Bayes* serta dataset yang digunakan diterapkan *split validation* dimana perbandingan data *training* dan data *testing* sebesar 80:20. Untuk melihat perbandingan akurasi sebelum dan sesudah ditambahkan *PSO* dapat kita lihat pada tabel dibawah ini.

Tabel 4. 6 Perbandingan *Naïve Bayes* dan *PSO + Naïve Bayes*

No	Keterangan	Naïve Bayes (80:20)	Naïve Bayes + Cross Validation (90:10)	PSO + Naïve Bayes (80:20)
1	<i>Accuracy</i>	87,60%	82,83%	89,26%
2	<i>Precision</i>	87,74%	82,10%	89,62%
3	<i>Recall / Confusion Matrix</i>	84,55%	79,17%	86,36%
4	<i>AUC</i>	0,933	0,900	0,933

Pada tabel diatas dapat dilihat bahwa dari percobaan menggunakan data *training* dan *testing* dengan perbandingan 80:20 dan 90:10 menghasilkan akurasi yang berbeda, dari penelitian ini dapat disimpulkan bahwa mengambil akurasi yang terbaik yaitu dengan data *training* 80%. Perbandingan hasil penelitian ini dapat dilihat pada grafik dibawah ini.

Perbandingan Accuracy Penelitian



Gambar 4. 24 Grafik Perbandingan *Accuracy* Penelitian

Pada hasil terdapat kenaikan akurasi pada metode dengan menerapkan *PSO* dan *Naïve Bayes* dengan pembagian dataset 80:20 sehingga pada penelitian ini dapat dibandingkan dengan penelitian sebelumnya. Perbandingan akurasi dari penelitian sebelumnya dapat dilihat pada tabel dibawah ini.

Tabel 4. 7 Perbandingan Akurasi Dengan Penelitian Sebelumnya

Metode	Algoritma	Akurasi
Peneliti sebelumnya : tanpa <i>optimasi feature selection</i> (El Hamdaoui, H., Boujraf, S., Chaoui, N. E. H., & Maaroufi, M., 2020)	<i>Naïve Bayes + Cross-Validation</i>	82,17%
	<i>Naïve Bayes + Split Data</i>	84,28%
Penelitian saat ini	<i>Naïve Bayes + Split Data</i>	87,60%
	<i>Naïve Bayes + Cross-Validation</i>	82,83%
	<i>PSO + Naïve Bayes</i>	89,26%

Dari tabel diatas bahwa terjadi peningkatan nilai akurasi pada penelitian saat ini, analisis dari peningkatan nilai akurasi ini dibandingkan dengan penelitian sebelumnya yaitu dari pengolahan data penelitian sehingga *record data* yang siap dimasukkan berbeda sehingga mempengaruhi nilai akurasi dimana akurasi penelitian sebelumnya yaitu 82.17% dengan *Naïve Bayes* dan 84,28% dengan *Split Data* , dan penelitian saat ini mengalami peningkatan menjadi 82,83% dengan *Naïve Bayes* dan 89.26% dengan *PSO* dan *Split Data*.