

## BAB IV

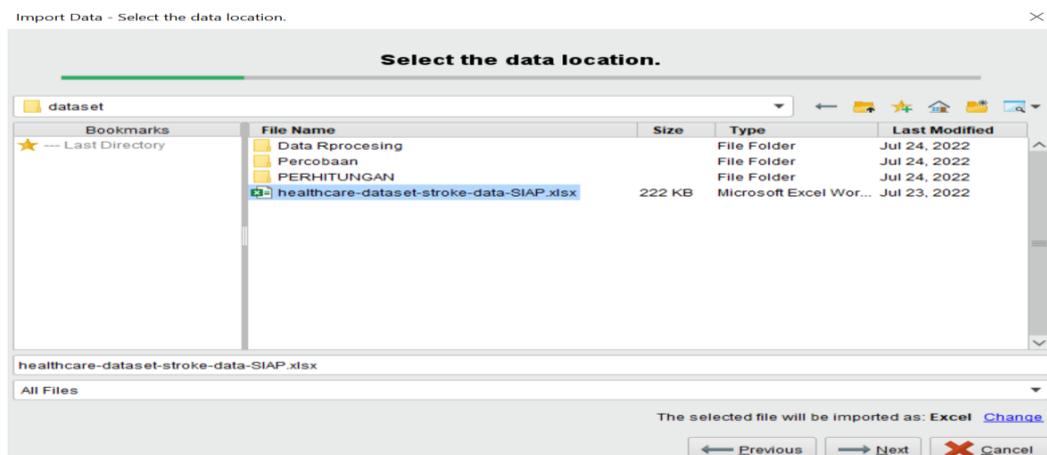
### HASIL DAN PEMBAHASAN

#### 4.1 Hasil Penelitian

Pada penelitian ini digunakan seleksi fitur PSO dan Algoritma *Decision Tree C4.5*. kasus yang diambil adalah data penyakit stroke yang berjumlah 5.110 data dan dilakukan *preprocessing* sehingga menjadi 3.426 data yang siap diolah. Data sample ini yang akan dihitung probabilitasnya, kemudian akan dicari tingkat akurasi tertinggi dari data tersebut dan akan diimplementasikan pada tools *RapidMiner*.

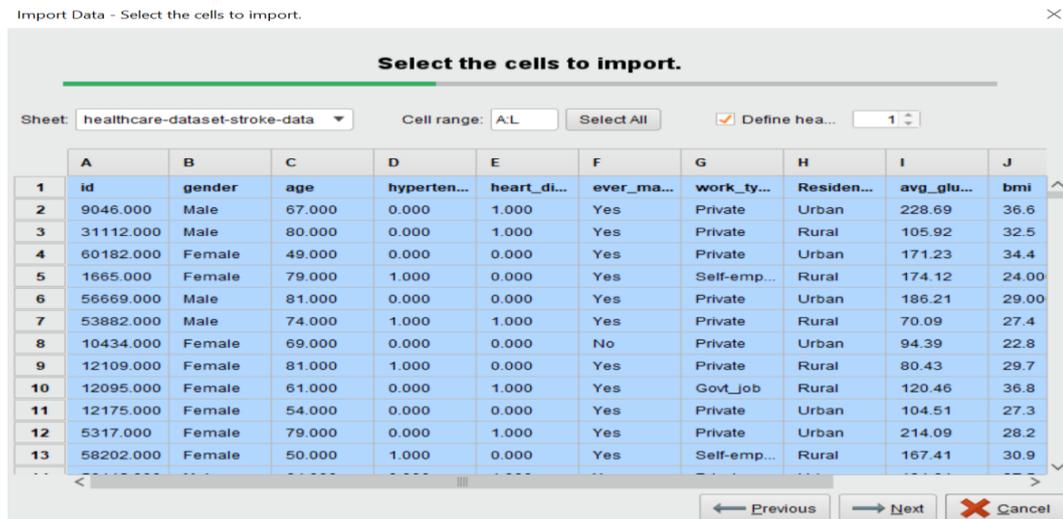
##### 4.1.1 Dataset

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*.



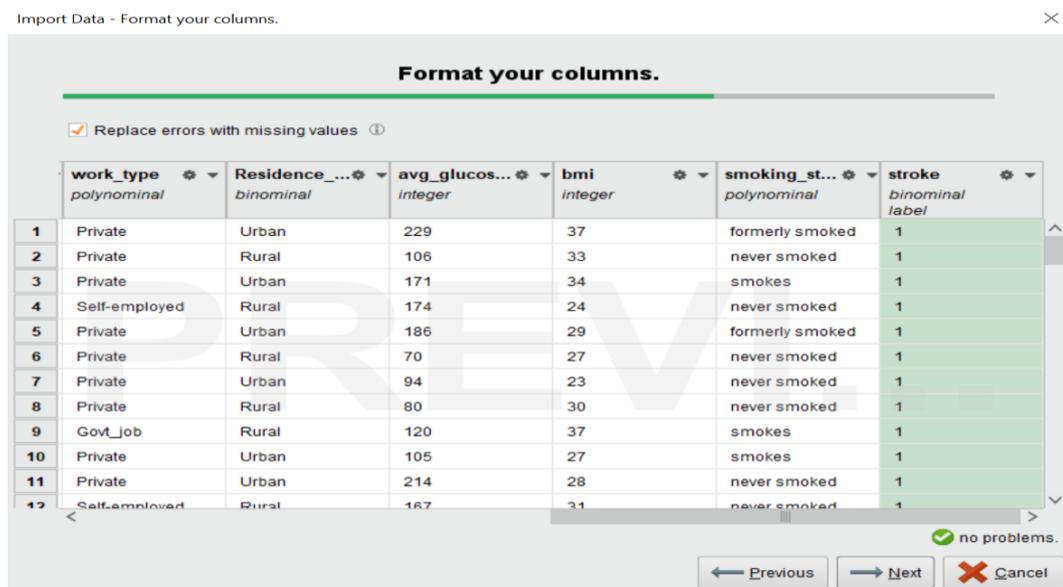
Gambar 4. 1 *Import Dataset*

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



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.

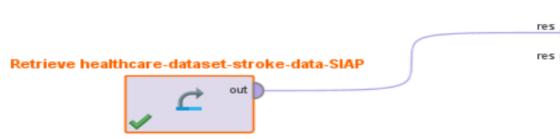


Gambar 4. 3 Format Type Dataset

Tabel 4. 1 Type Data Atribut

No	Atribut	Type Data
1	id	Integer
2	Gender	Binominal
3	Age	Integer
4	Hypertension	Binominal
5	Heart_disease	Binominal
6	Ever_married	Binominal
7	Work_type	Polynomial
8	Residence_type	Binominal
9	Avg_glucose_level	Integer
10	Bmi	Integer
11	Smoking_status	Polynomial
12	Stroke	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.

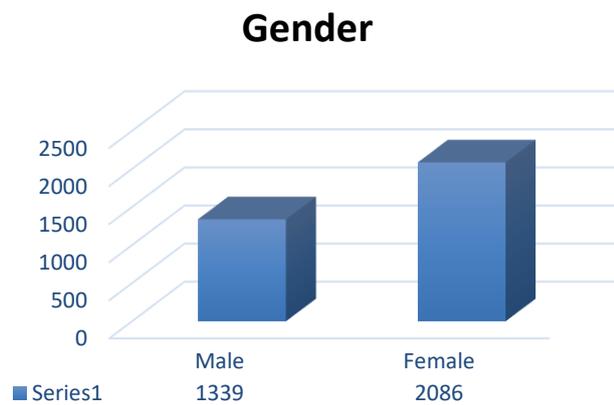
Name	Type	Missing	Statistics	Filter (12 / 12 attributes):
id	Integer	0	Min: 84, Max: 72915, Average: 37333.512	Search for Attributes
stroke	Binominal	0	Least: 1 (180), Most: 0 (3245), Values: 0 (3245), 1 (180)	
gender	Binominal	0	Least: Male (1339), Most: Female (2086), Values: Female (2086), Male (1339)	
age	Integer	0	Min: 10, Max: 82, Average: 48.653	
hypertension	Binominal	0	Least: 1 (408), Most: 0 (3017), Values: 0 (3017), 1 (408)	
heart_disease	Binominal	0	Least: 1 (206), Most: 0 (3219), Values: 0 (3219), 1 (206)	
ever_married	Binominal	0	Least: No (826), Most: Yes (2599), Values: Yes (2599), No (826)	

Gambar 4. 5 Statistik Dataset

#### 4.1.2 Model Data Masukkan

##### a. Gender

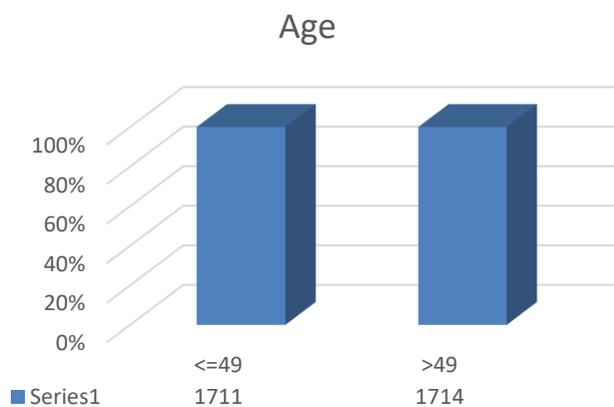
Gambar grafik dibawah merupakan gambaran karakteristik data *gender* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 1339 *male* dan 2086 *female*.



Gambar 4. 6 Data *Gender*

##### b. Age

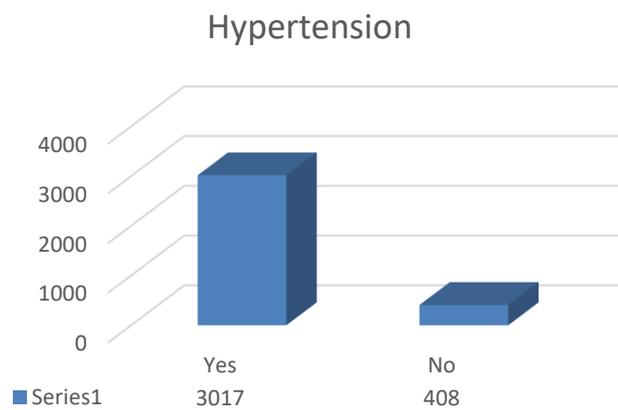
Gambar grafik dibawah merupakan gambaran karakteristik data *age* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 1711 *age*  $\leq 49$  dan 1714 *age*  $> 49$ .



Gambar 4. 7 Data *Age*

c. *Hypertension*

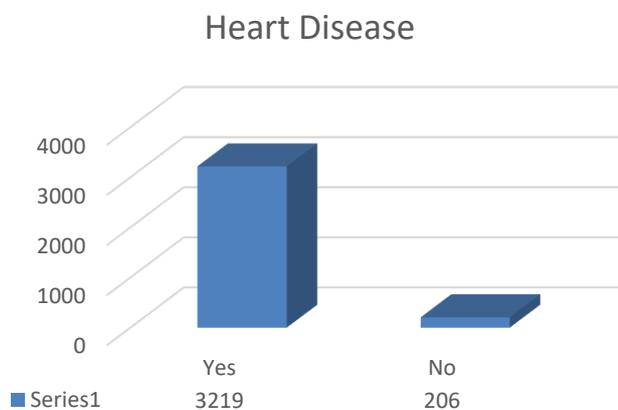
Gambar grafik dibawah merupakan gambaran karakteristik data *hypertension* terdiri dari 1=Yes dan 0=No dari keseluruhan dataset sebanyak 3425 record, digambarkan terdapat 3017 terkena *hypertension* dan 408 tidak terkena *hypertension*.



Gambar 4. 8 Data *Hypertension*

d. *Heart Disease*

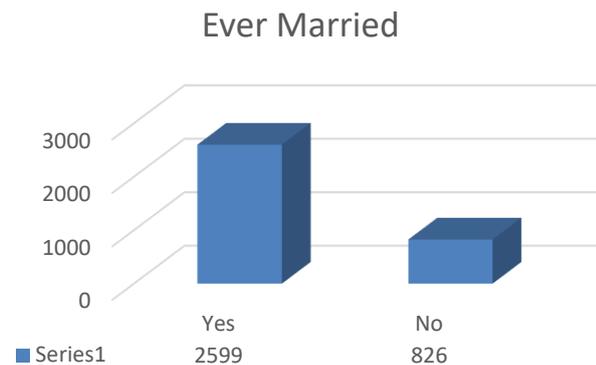
Gambar grafik dibawah merupakan gambaran karakteristik data *heart disease* terdiri dari 1=Yes dan 0=No dari keseluruhan dataset sebanyak 3425 record, digambarkan terdapat 3219 terkena *heart disease* dan 206 tidak terkena *heart disease*.



Gambar 4. 9 Data *Heart Disease*

e. *Ever Married*

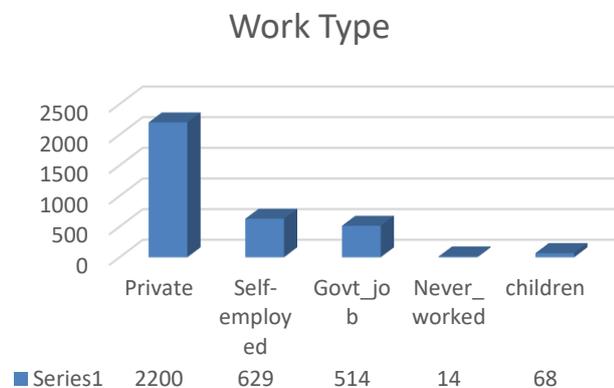
Gambar grafik dibawah merupakan gambaran karakteristik data *heart disease* terdiri dari *Yes* dan *No* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 2599 dengan status *Yes* dan 826 dengan status *No*.



Gambar 4. 10 Data *Ever Married*

f. *Work Type*

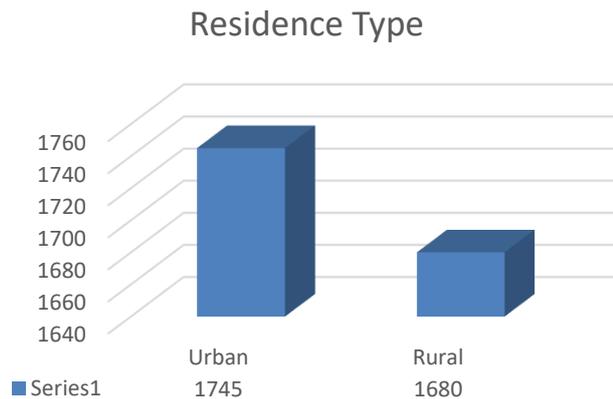
Gambar grafik dibawah merupakan gambaran karakteristik data *work type* terdiri dari *private*, *self employe*, *govt job*, *never worked*, *children* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 2200 data *private*, 629 data *self employed*, 514 data *govt job*, 14 data *never worked* dan 68 data *children*.



Gambar 4. 11 Data *Work Type*

g. *Residence Type*

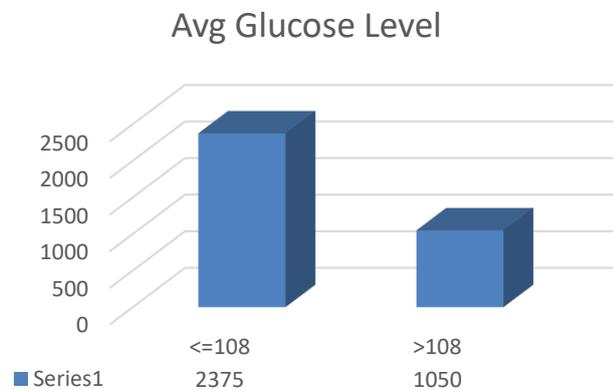
Gambar grafik dibawah merupakan gambaran karakteristik data *residence type* terdiri dari *urban* dan *rural* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 1745 data *urban*, 1680 data *rural*.



Gambar 4. 12 Data *Residence Type*

h. *Avg Glucose Level*

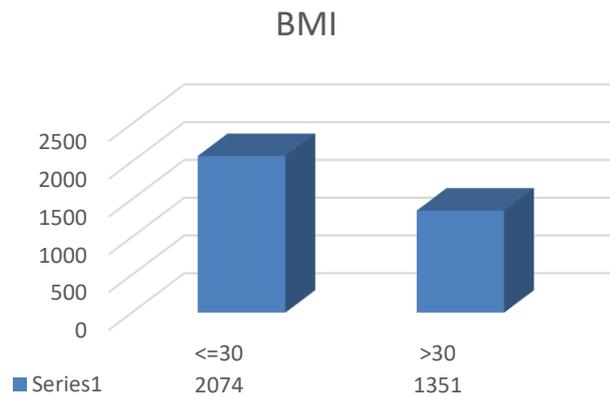
Gambar grafik dibawah merupakan gambaran karakteristik data *avg glucose level* terdiri dari  $\leq 108$  dan  $> 108$  dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 2375 data  $\leq 108$ , 1050 data  $> 108$ .



Gambar 4. 13 Data *Avg Glucose Level*

i. *BMI*

Gambar grafik dibawah merupakan gambaran karakteristik data *bmi* terdiri dari  $\leq 30$  dan  $> 30$  dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 2074 data  $\leq 30$ , 1351 data  $> 30$ .

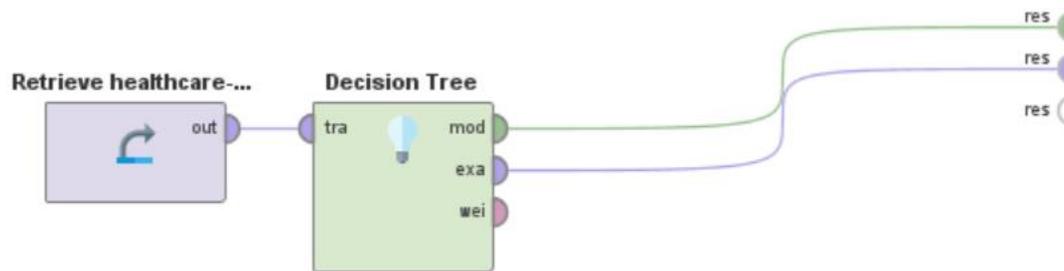
Gambar 4. 14 Data *BMI*j. *Smoking Status*

Gambar grafik dibawah merupakan gambaran karakteristik data *smoking status* terdiri dari *formerly smoked*, *never smoked* dan *smokes* dari keseluruhan dataset sebanyak 3425 *record*, digambarkan terdapat 836 data *formerly smoked*, 1852 data *never smoked* dan 737 *smokes*.

Gambar 4. 15 Data *Smoking Status*

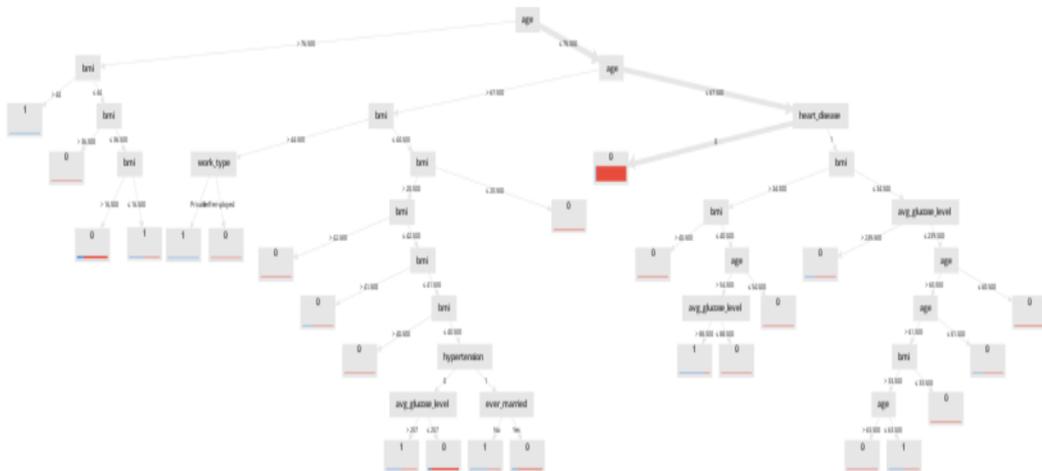
### 4.1.3 Algoritma *Decision Tree* C4.5

Pada tahap ini yaitu dataset yang sudah di *import* kedalam rapidminer akan dilakukan klasifikasi dengan menggunakan model algoritma *Decision Tree* C4.5 untuk menentukan pola dan hasil dari pohon keputusan.



Gambar 4. 16 Tampilan Proses *Decision Tree* C4.5

Berikut ini adalah tampilan pohon keputusan pada algoritma *Decision Tree* C4.5 dimana *node root* awal yaitu *age*.



Gambar 4. 17 Tampilan Pohon Keputusan *Decision Tree* C4.5



Mencari entropy dan gain untuk menentukan node root awal, diketahui dari proses diatas bahwa dengan menggunakan tools rapidminer node root awal adalah age. Disini kita akan melakukan pengukuran nilai entropy dan gain dengan cara menghitung manual. Proses ini adalah untuk pembuktian apakah tools rapidminer berjalan dengan baik.

Menghitung nilai entropy total pada dataset

		Jumlah (S)	Yes (Si)	No (Si)	Entropy	Gain
<b>Total</b>		3425	180	3245	0,297152	

Rumus :

$$\text{Entropy (S)} \sum_{i=1}^n -p_i \log_2 p_i$$

$$= -\left(\left(\frac{180}{3425}\right) \times \log_2 \left(\frac{180}{3425}\right)\right) + \left(\left(\frac{3245}{3425}\right) \times \log_2 \left(\frac{3245}{3425}\right)\right) = 0,297152$$

Setelah diketahui entropy total pada dataset, maka kita selanjutnya akan menghitung entropy dan gain pada setiap atribut.

Menghitung entropy dan gain pada atribut gender

		Jumlah (S)	Yes (Si)	No (Si)	Entropy	Gain
<b>Total</b>		3425	180	3245	0,297152	
<b>gender</b>						<b>0,000110379</b>
	Male	1339	75	1264	0,311406	
	Female	2086	105	1981	0,287821	

$E(\text{semua (Male)})$

$$= -\left(\left(\frac{75}{1339}\right) \times \log_2 \left(\frac{75}{1339}\right)\right) + \left(\left(\frac{1264}{1339}\right) \times \log_2 \left(\frac{1264}{1339}\right)\right) = 0,311406$$

$E$  (semua (Female))

$$= - \left( \left( \frac{105}{2086} \right) \times \log_2 \left( \frac{105}{2086} \right) \right) + \left( \left( \frac{1981}{2086} \right) \times \log_2 \left( \frac{1981}{2086} \right) \right) = 0,287821$$

$G$  (semua (Gender))

$$= - \left( \left( \frac{1339}{3425} \right) \times 0,311406 \right) + \left( \left( \frac{2086}{3425} \right) \times 0,287821 \right) = 0,000110379$$

$Sinfo$  (semua (Gender))

$$= \left( - \left( \frac{1339}{3425} \right) \times \log_2 \left( \frac{1339}{3425} \right) - \left( \frac{2086}{3425} \right) \times \log_2 \left( \frac{2086}{3425} \right) \right) = 0,965409181$$

$Gratio$  (semua (Gender))

$$= \frac{0,000110379}{0,965409181} = 0,000114334$$

Lakukan perhitungan diatas ke seluruh atribut, maka setelah dilakukan perhitungan

didasar diatas didapatkan hasil entrophy dan gain setiap atribut seperti dibawah ini.

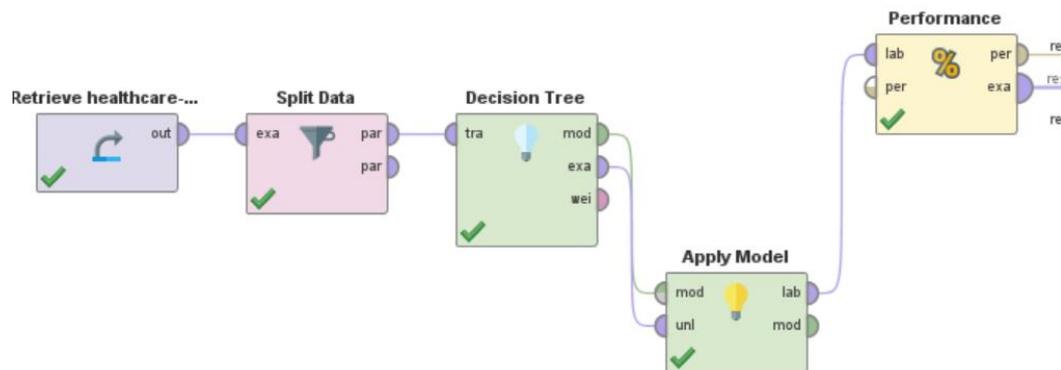
Algoritma DTC4.5								
		Jumlah (S)	Yes (Si)	No (Si)	Entropy	Gain	Split Info	Gain Ratio
<b>Total</b>		3425	180	3245	0,297152			
<b>gender</b>						0,00011	0,96541	0,00011
	Male	1339	75	1264	0,311406			
	Female	2086	105	1981	0,287821			
<b>Age</b>						0,03221	1,00000	0,03221
	<=49	1711	15	1696	0,072502			
	>49	1714	165	1549	0,457046			
<b>hypertension</b>						0,01112	0,52684	0,02111
	0	3017	123	2894	0,245807			
	1	408	57	351	0,583468			
<b>heart_disease</b>						0,00921	0,32802	0,02807

	0	3219	144	3075	0,263593			
	1	206	36	170	0,668473			
<b>ever_married</b>						0,00433	0,79698	0,00543
	Yes	2599	160	2439	0,333615			
	No	826	20	806	0,164483			
<b>work_type</b>						0,00342	1,41455	0,00241
	Private	2200	109	2091	0,284463			
	Self-employed	629	48	581	0,389048			
	Govt_job	514	23	491	0,263649			
	Never_worked	14	0	14	0			
	children	68	0	68	0			
<b>Residence_type</b>						0,00003	0,99974	0,00003
	Urban	1745	94	1651	0,302607			
	Rural	1680	86	1594	0,291433			
<b>avg_glucose_level</b>						0,00689	0,88917	0,00775
	<=108	2375	89	2286	0,230587			
	>108	1050	91	959	0,425245			
<b>bmi</b>						0,00013	0,96761	0,00013
	<=30	2074	104	1970	0,287011			
	>30	1351	76	1275	0,312394			
<b>smoking_status</b>						0,00122	1,45316	0,00084
	formerly smoked	836	57	779	0,359102			
	never smoked	1852	84	1768	0,266334			
	smokes	737	39	698	0,298662			
<b>Nilai Gain Ratio Maksimal</b>								<b>0,03221</b>

Maka didapatkan hasil gain ratio terbesar yaitu 0,032209154 dimana attribute tersebut adalah age, atribut ini akan menjadi node root awal sebagai pohon keputusan, selanjutnya untuk mencari cabang node yang lain lakukan Kembali pencarian entropy, gain, split info dan gain ratio pada seluruh attribute.

#### 4.1.4 Algoritma *Decision Tree* C4.5 Dengan Split Validation

Penerapan algoritma *Decision Tree* 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. 19 Proses *Decision Tree* C4.5 dan *Split Validation*

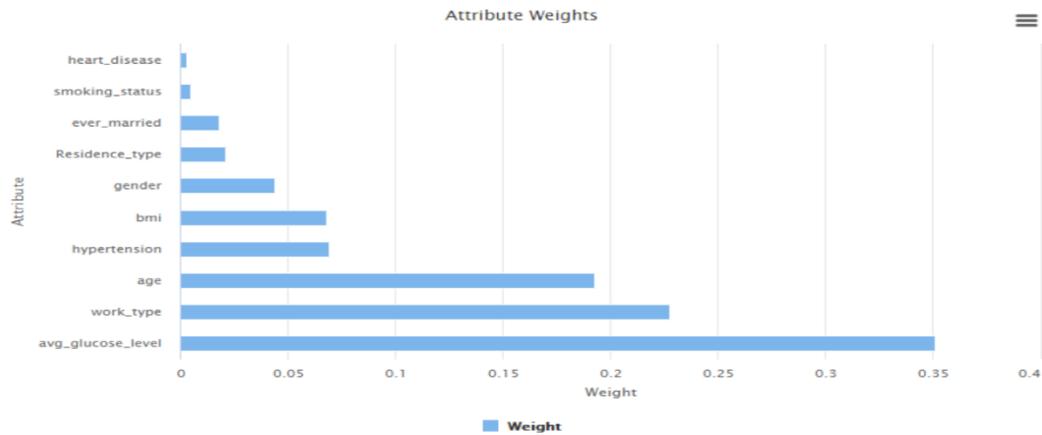
Atribut yang digunakan pada algoritma *Decision Tree* C4.5 ini melibatkan semua atribut yang berjumlah 10, Adapun atribut dan bobot dapat dilihat dari tabel berikut.

Tabel 4. 2 Tabel Atribut Yang Digunakan Sebelum PSO

NO	Attribute	Weight
1	smoking	0.005
2	gender	0.044
3	Work_type	0.228
4	Residence	0.021
5	Heart_disease	0.003
6	Avg_glucose_lvl	0.351
7	Hypertension	0.096
8	Age	0.193
9	Bmi	0.068
10	Ever_married	0.018

Dari keterangan tabel diatas bahwa setiap atribut mempunyai pembobotan masing masing untuk digunakan sebagai model klasifikasi prediksi stroke dengan menggunakan algoritma *Decision Tree* C4.5.

Dari pembobotan atribut diatas maka dapat dilihat juga visualisasi grafik perbandingan antar atribut seperti gambar dibawah ini.



Gambar 4. 20 Gambar Pembobotan Atribut Dengan Algoritma Decision Tree C4.5

Hasil Akurasi *Decision Tree C4.5* dan *Split Validation*, pada pembagian data sampling 80% dan data testing 20% maka didapatkan akurasi sebesar 95,29%

accuracy: 95.29%

	true 1	true 0	class precision
pred. 1	30	15	66.67%
pred. 0	114	2581	95.77%
class recall	20.83%	99.42%	

Gambar 4. 21 Nilai Akurasi Dengan Split Validation

Menghitung akurasi :

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

$$= \frac{30 + 2581}{30 + 2581 + 15 + 114} = \frac{2611}{2740} = 95,29\%$$

Berikut adalah hasil *precision* yaitu mendapatkan 95,77% seperti gambar dibawah ini.

precision: 95.77% (positive class: 0)

	true 1	true 0	class precision
pred. 1	30	15	66.67%
pred. 0	114	2581	95.77%
class recall	20.83%	99.42%	

Gambar 4. 22 Nilai Precision Dengan Split Validation

Menghitung Precision, Rumus :  $Precision = \frac{TP}{TP+FP}$

$$P(1) = \frac{30}{30+15} = \frac{30}{45} = 66.67\%$$

$$P(0) = \frac{2581}{2581+114} = \frac{2581}{2701} = 95.77\%$$

Berikut adalah hasil *confusion matrix* yaitu mendapatkan 99,42% seperti gambar dibawah ini.

recall: 99.42% (positive class: 0)

	true 1	true 0	class precision
pred. 1	30	15	66.67%
pred. 0	114	2581	95.77%
class recall	20.83%	99.42%	

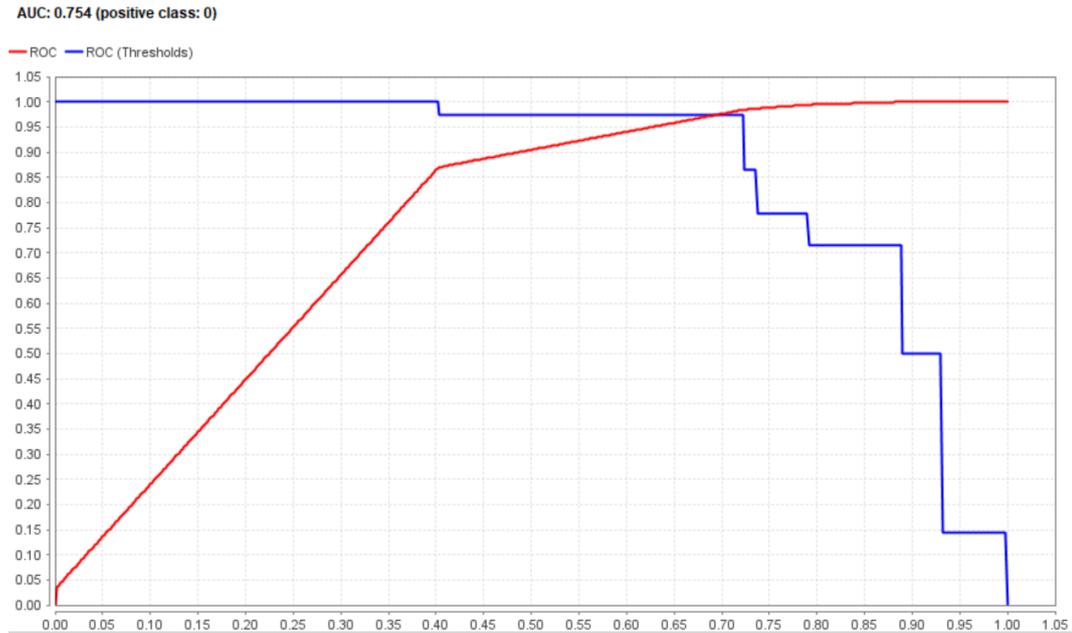
Gambar 4. 23 Nilai Confusion Matrix Dengan Split Validation

Menghitung Recall, Rumus :  $Recall = \frac{TP}{TP+FN}$

$$R(1) = \frac{30}{30+114} = \frac{30}{144} = 20,83\%$$

$$R(0) = \frac{2581}{2581+15} = \frac{2581}{2596} = 99,42\%$$

Berikut adalah hasil kurva AUC yaitu mendapatkan nilai 0,754 seperti gambar dibawah ini.



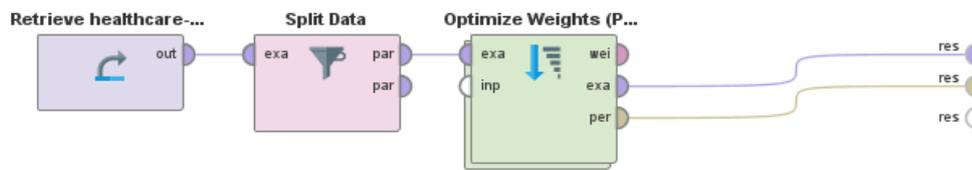
Gambar 4. 24 Kurva AUC Dengan Split Validation

Tabel 4. 3 Hasil dari *Decision Tree* dan *Split Validation* (80:20)

No	DTC4.5 (80:20)	Hasil Nilai
1	<i>Accuracy</i>	95.29%
2	<i>Precision</i>	95.77%
3	<i>Recall / Confusion Matrix</i>	99.42%
4	AUC	0.754

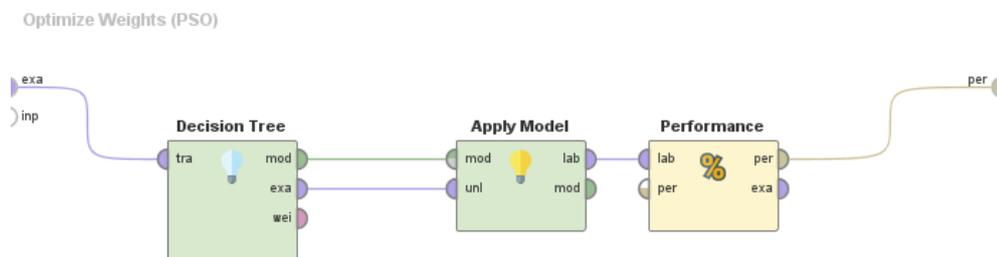
#### 4.1.5 *Patricle Swarm Optimization* (PSO) dan *Decision Tree* C4.5 Dengan *Split Data* (80:20)

Metode ini adalah untuk meningkatkan akurasi dalam klasifikasi dengan menggunakan Teknik seleksi fitur pembobotan atribut dengan menggunakan PSO dengan *Decision Tree* C4.5 serta perbandingan split data training sebesar 80% dan testing 20%. penerapan metode ini dapat dilihat pada gambar dibawah ini.



Gambar 4. 25 Proses PSO dan *Decision Tree* C4.5

Setelah seperti gambar diatas kita masukkan algoritma *Decision Tree* C4.5 kedalam PSO seperti gambar dibawah ini.



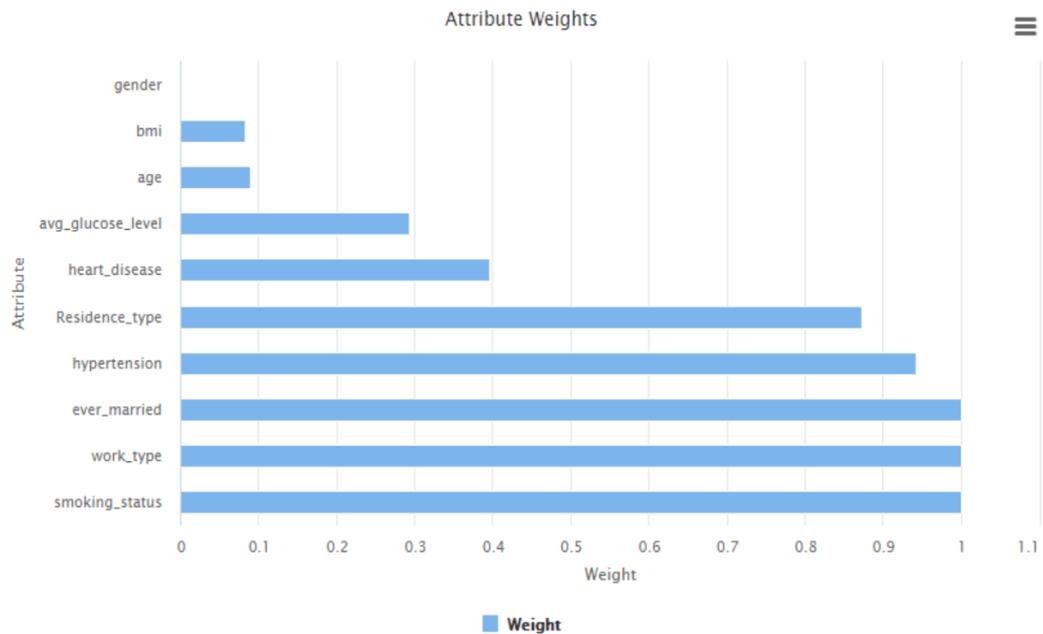
Gambar 4. 26 Sub Proses PSO dan *Decision Tree* C4.5

Atribut yang digunakan pada algoritma *Decision Tree* C4.5 ini melibatkan semua atribut yang berjumlah 10, Adapun atribut dan bobot dapat dilihat dari tabel berikut.

Tabel 4. 4 Tabel Atribut Yang Digunakan Sesudah PSO

NO	Attribute	Weight
1	smoking	1
2	gender	0
3	Work_type	1
4	Residence	0.873
5	Heart_disease	0.396
6	Avg_glucose_lvl	0.294
7	Hypertension	0.943
8	Age	0.089
9	Bmi	0.083
10	Ever_married	1

Dari keterangan tabel diatas bahwa terdapat atribut yang memiliki nilai 0 yaitu gender yang artinya pada proses PSO atribut tersebut tidak berkontribusi dalam proses klasifikasi prediksi stroke dengan menggunakan algoritma *Decision Tree* C4.5.



Gambar 4. 27 Gambar Pembobotan Atribut Dengan PSO

Dari pengaplikasian PSO dan *Decision Tree* C4.5 pada rapidminer maka dihasilkan peningkatan akurasi, dimana dapat dilihat perbandingan akurasi, precision dan recall serta AUC sesudah dan sebelum di tambahkan metode PSO.

Berikut adalah hasil *accuracy* yaitu mendapatkan 96,20% seperti gambar dibawah ini.

accuracy: 96.20%

	true 1	true 0	class precision
pred. 1	54	14	79.41%
pred. 0	90	2582	96.63%
class recall	37.50%	99.46%	

Gambar 4. 28 Nilai Akurasi PSO dan *Decision Tree* C4.5

Berikut adalah hasil *precision* yaitu mendapatkan 96,63% seperti gambar dibawah ini.

precision: 96.63% (positive class: 0)

	true 1	true 0	class precision
pred. 1	54	14	79.41%
pred. 0	90	2582	96.63%
class recall	37.50%	99.46%	

Gambar 4. 29 Nilai *Precision* PSO dan *Decision Tree* C4.5

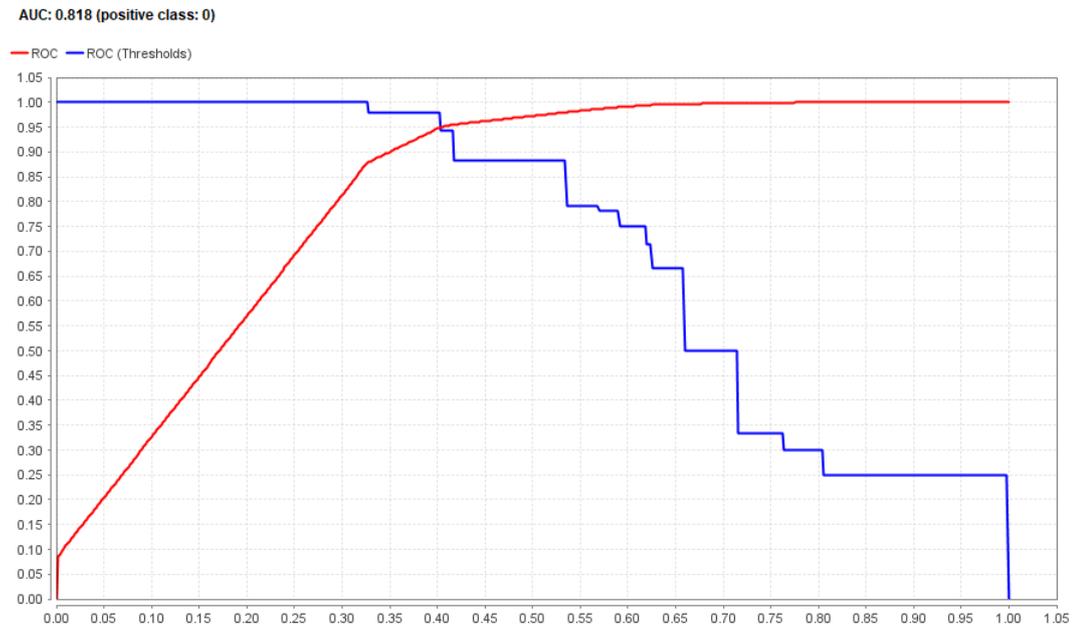
Berikut adalah hasil *confusion matrix* yaitu mendapatkan 99,46% seperti gambar dibawah ini.

recall: 99.46% (positive class: 0)

	true 1	true 0	class precision
pred. 1	54	14	79.41%
pred. 0	90	2582	96.63%
class recall	37.50%	99.46%	

Gambar 4. 30 Nilai *Confusion Matrix* PSO dan *Decision Tree* C4.5

Berikut adalah hasil kurva AUC yaitu mendapatkan nilai 0,818 seperti gambar dibawah ini.



Gambar 4. 31 Kurva AUC PSO dan *Decision Tree C4.5*

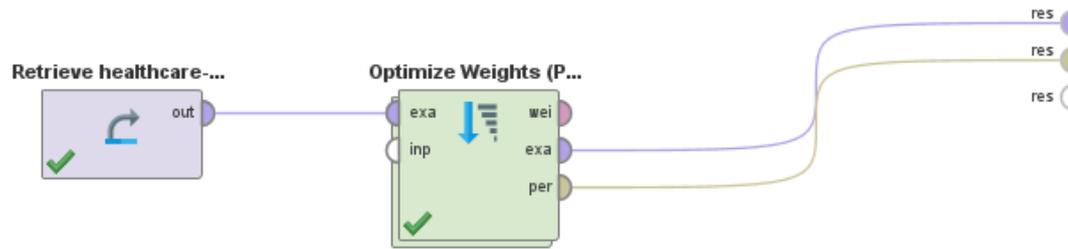
Tabel 4. 5 Hasil dari PSO dan *Decision Tree* (80:20)

No	PSO + DTC4.5 (80:20)	Hasil Nilai
1	<i>Accuracy</i>	96.20%
2	<i>Precision</i>	96.63%
3	<i>Recall / Confusion Matrix</i>	99.46%
4	AUC	0.818

#### 4.1.6 *Particle Swarm Optimization (PSO) dan Decision Tree C4.5*

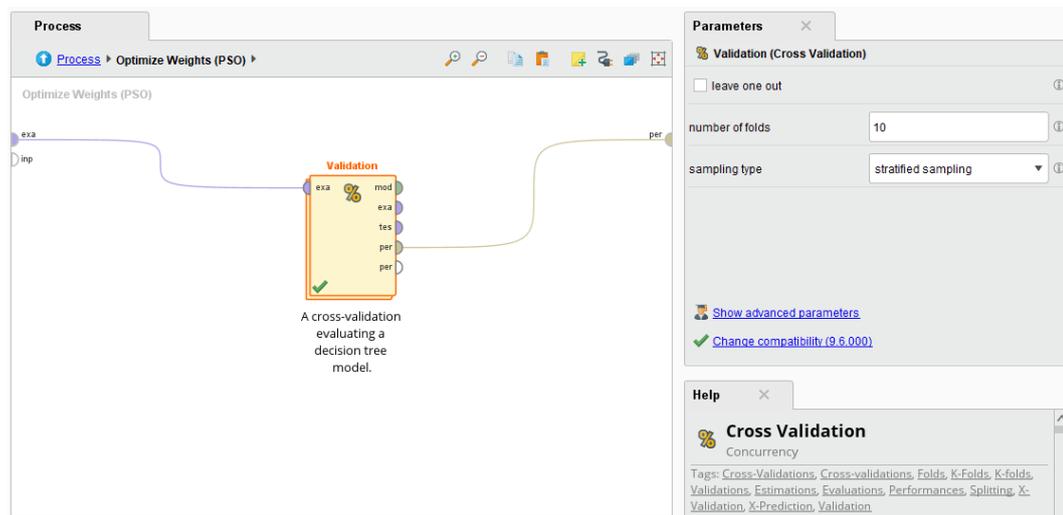
##### **Dengan *Cross Validation* (90:10)**

Metode ini adalah untuk mengetahui dan membandingkan tingkat akurasi dalam klasifikasi dengan menggunakan *cross validation* yang di implementasikan dengan menggunakan PSO dengan *Decision Tree C4.5*. untuk perbandingan data training sebesar 90% dan testing 10%. penerapan metode ini dapat dilihat pada gambar dibawah ini.



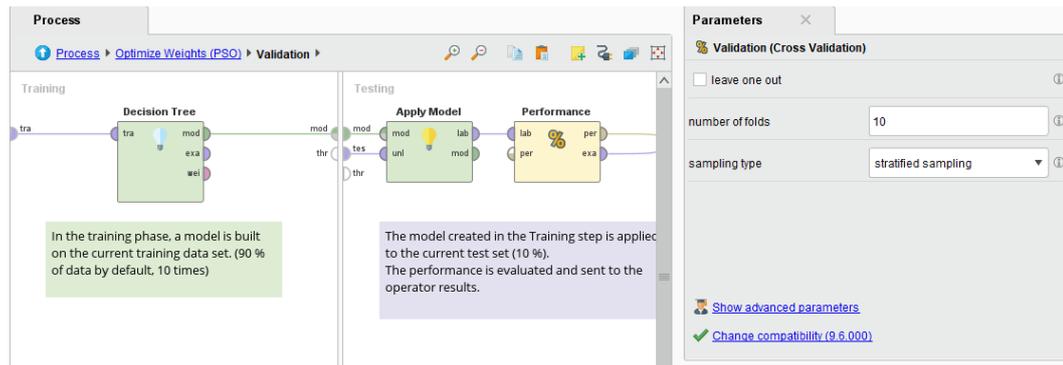
Gambar 4. 32 Proses PSO dan *Decision Tree* C4.5 Dengan *Cross Validation*

Setelah seperti gambar diatas kita masukkan *cross validation* pada sub proses di PSO dengan *number of fold* 10.



Gambar 4. 33 Sub Proses PSO dan *Decision Tree* C4.5

Selanjutnya masukkan algoritma dengan data training dan testing 90:10 dengan *number of folds* 10 seperti gambar dibawah ini.



Gambar 4. 34 Proses PSO Validation

Sehingga didapatkan hasil *accuracy* yaitu mendapatkan 94,77% seperti gambar dibawah ini.

accuracy: 94.77% +/- 0.21% (micro average: 94.77%)

	true 1	true 0	class precision
pred. 1	1	0	100.00%
pred. 0	179	3245	94.77%
class recall	0.56%	100.00%	

Gambar 4. 35 Nilai Akurasi PSO dan *Decision Tree* C4.5 (Cross Validation)

Berikut adalah hasil *precision* yaitu mendapatkan 94,77% seperti gambar dibawah ini.

precision: 94.77% +/- 0.21% (micro average: 94.77%) (positive class: 0)

	true 1	true 0	class precision
pred. 1	1	0	100.00%
pred. 0	179	3245	94.77%
class recall	0.56%	100.00%	

Gambar 4. 36 Nilai Precision PSO dan *Decision Tree* C4.5 (Cross Validation)

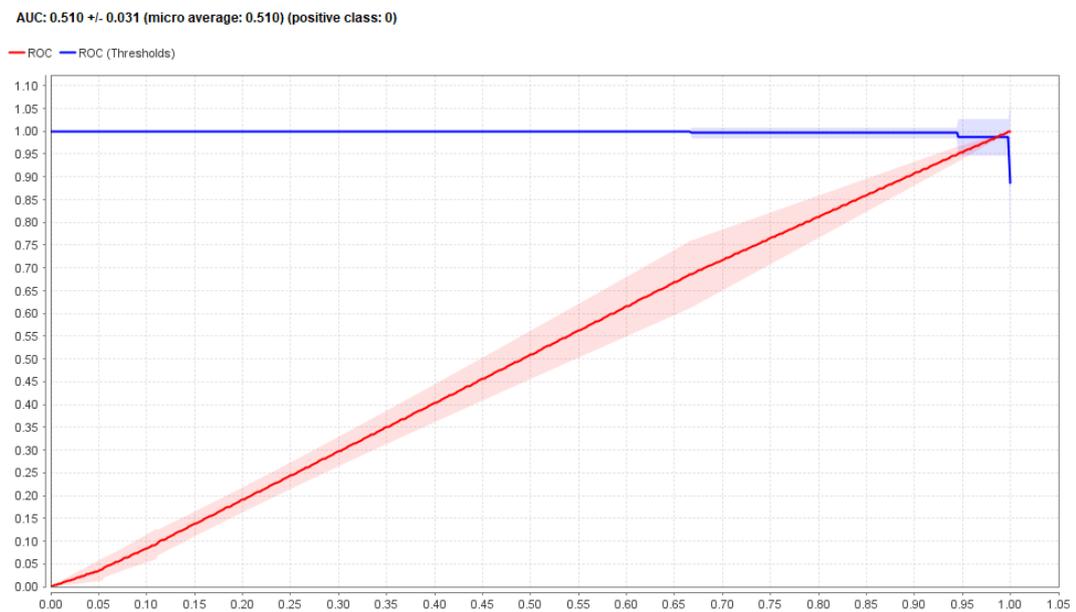
Berikut adalah hasil *confusion matrix* yaitu mendapatkan 100% seperti gambar dibawah ini.

recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: 0)

	true 1	true 0	class precision
pred. 1	1	0	100.00%
pred. 0	179	3245	94.77%
class recall	0.56%	100.00%	

Gambar 4. 37 Nilai Confusion Matrix PSO dan *Decision Tree* C4.5 (Cross Validation)

Berikut adalah hasil kurva AUC yaitu mendapatkan nilai 0,510 seperti gambar dibawah ini.



Gambar 4. 38 Kurva AUC PSO dan *Decision Tree* C4.5 (Cross Validation)

Tabel 4. 6 Hasil dari PSO dan *Decision Tree* Cross Validation (90:10)

No	PSO + DTC4.5 (80:20)	Hasil Nilai
1	<i>Accuracy</i>	94.77%
2	<i>Precision</i>	94.77%
3	<i>Recall / Confusion Matrix</i>	100%
4	AUC	0.510

## 4.2 Pembahasan Hasil Penelitian

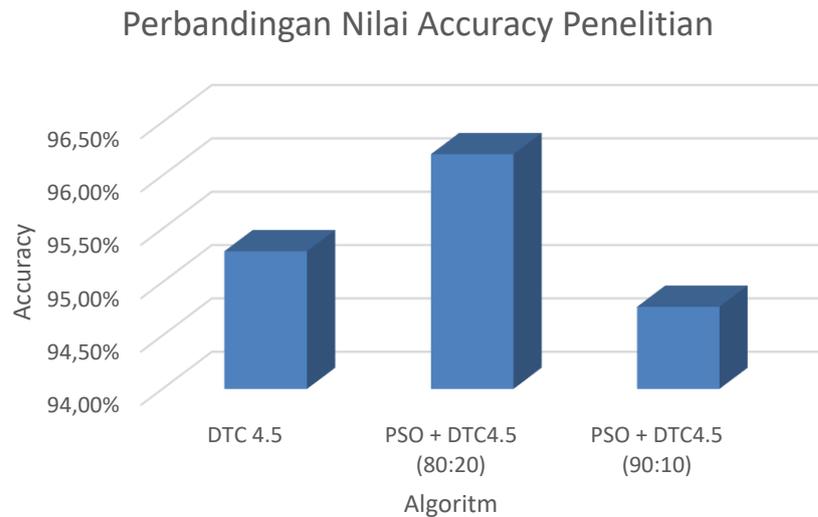
Pembahasan hasil penelitian yaitu membahas hasil penelitian yang telah dilakukan. Penelitian yang dilakukan yang pertama yaitu mengklasifikasi penyakit stroke dengan menggunakan algoritma *decision tree* C4.5 dan penelitian ke dua yaitu untuk menabah akurasi sehingga ditambahkan metode seleksi fitur PSO dan Algoritma *Decision tree* C4.5 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. 7 Perbandingan DTC4.5 dan PSO+DCT5.4

No	Keterangan	DTC4.5	PSO + DTC4.5 (80:20)	PSO + DTC4.5 (90:10)
1	<i>Accuracy</i>	95.29%	<b>96.20%</b>	94.77%
2	<i>Precision</i>	95.77%	<b>96.63%</b>	94.77%
3	<i>Recall / Confusion Matrix</i>	99.42%	<b>99.46%</b>	100%
4	AUC	0.754	<b>0.818</b>	0.510

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.



Gambar 4. 39 Gambar Grafik Perbandingan *Accuracy* Penelitian

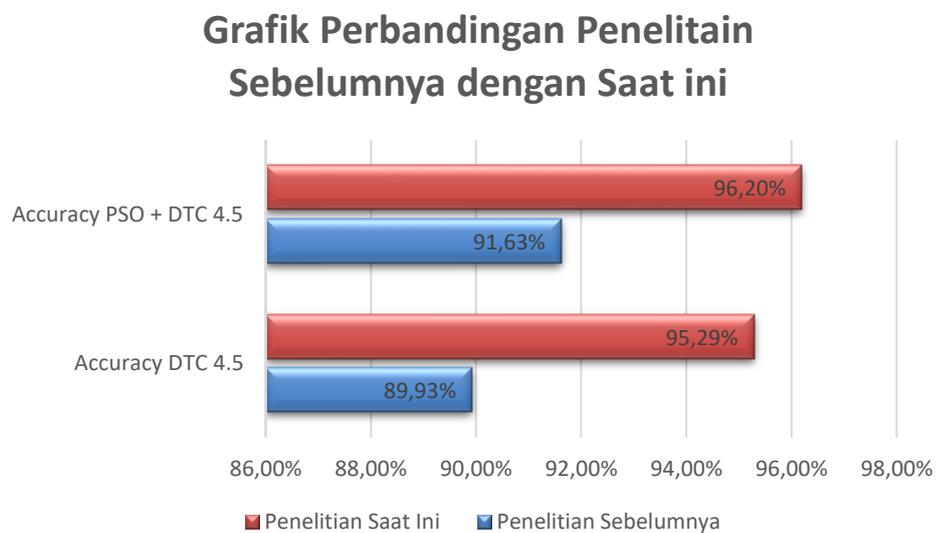
Pada hasil terdapat kenaikan akurasi pada metode dengan menerapkan PSO dan DTC4.5 dengan pembagian dataset 80:20 sehingga pada penelitian ini dapat dibandingkan dengan penelitian sebelumnya. Perbandingan akurasi dari penelitian sebelumnya dapat dilihat pada tabel dibawah ini.

Tabel 4. 8 Perbandingan Akurasi Dengan Penelitian Sebelumnya

Penelitian	Algoritma	Akurasi	AUC
Penelitian Sebelumnya, Ramdhan Saepul Rohman 2020	C4.5	89,93%	0.610
	C4.5 + PSO	91,63%	0.500
<b>Penelitian saat ini</b>	<b>C4.5</b>	<b>95,29%</b>	<b>0.754</b>
	<b>C4.5 + PSO</b>	<b>96.20%</b>	<b>0.818</b>

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 di

masukkan berbeda sehingga mempengaruhi nilai akurasi dimana akurasi penelitian sebelumnya yaitu 89,93% dengan C4.5 dan 91,63% dengan PSO mengalami peningkatan menjadi 95,29% dengan C4.5 dan 96,20% dengan PSO, serta nilai AUC juga mengalami peningkatan seperti tabel diatas. Adapun grafik perbandingan untuk melihat perbandingan penelitian sebelumnya dan saat ini bisa dilihat dibawah ini.



*Gambar 4. 40 Gambar Grafik Perbandingan Hasil Penelitian Sebelumnya  
Dengan Saat Ini.*