

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Bussiness Understanding atau Pemahaman Bisnis

Tujuan dari penelitian ini adalah untuk menghasilkan model algoritma yang memiliki akurasi sangat tinggi menggunakan metode seleksi fitur Particle Swam Optimization (PSO) dan algoritma Decision Tree C4.5.

#### 4.2 Data Understanding atau Pemahaman Data

Dataset yang digunakan pada penelitian ini berasal dari situs UCI Machine Learning.

Tabel 4.1 Class Label

Class Label	Data Training	Data Testing
Y (Deposito)	70%	30%

#### 4.3 Data Preparation atau Persiapan Data

Dalam penelitian ini, data yang ada pada penelitian ini adalah data sekunder dimana data sudah tersedia dari situs web uci machine learning sehingga tidak ada pengambilan data langsung kepada perusahaan. Situs pengambilan data yang digunakan yaitu <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>. Atribut-atribut tersebut dijelaskan pada table 4.2 dibawah ini:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
2	30	unemploy	married	primary	no	1787	no	no	cellular	19	oct	79	1	-1	0	unknown	no
3	33	services	married	secondary	no	4789	yes	yes	cellular	11	may	220	1	339	4	failure	no
4	35	managemt	single	tertiary	no	1350	yes	no	cellular	16	apr	185	1	330	1	failure	no
5	30	managemt	married	tertiary	no	1476	yes	yes	unknown	3	jun	199	4	-1	0	unknown	no
6	59	blue-collar	married	secondary	no	0	yes	no	unknown	5	may	226	1	-1	0	unknown	no
7	35	managemt	single	tertiary	no	747	no	no	cellular	23	feb	141	2	176	3	failure	no
8	36	self-emplo	married	tertiary	no	307	yes	no	cellular	14	may	341	1	330	2	other	no
9	39	technician	married	secondary	no	147	yes	no	cellular	6	may	151	2	-1	0	unknown	no
10	41	entrepreni	married	tertiary	no	221	yes	no	unknown	14	may	57	2	-1	0	unknown	no
11	43	services	married	primary	no	-88	yes	yes	cellular	17	apr	313	1	147	2	failure	no
12	39	services	married	secondary	no	9374	yes	no	unknown	20	may	273	1	-1	0	unknown	no
13	43	admin.	married	secondary	no	264	yes	no	cellular	17	apr	113	2	-1	0	unknown	no
14	36	technician	married	tertiary	no	1109	no	no	cellular	13	aug	328	2	-1	0	unknown	no
15	20	student	single	secondary	no	502	no	no	cellular	30	apr	261	1	-1	0	unknown	yes
16	31	blue-collar	married	secondary	no	360	yes	yes	cellular	29	jan	89	1	241	1	failure	no
17	40	managemt	married	tertiary	no	194	no	yes	cellular	29	aug	189	2	-1	0	unknown	no
18	56	technician	married	secondary	no	4073	no	no	cellular	27	aug	239	5	-1	0	unknown	no
19	37	admin.	single	tertiary	no	2317	yes	no	cellular	20	apr	114	1	152	2	failure	no
20	25	blue-collar	single	primary	no	-221	yes	no	unknown	23	may	250	1	-1	0	unknown	no
21	31	services	married	secondary	no	132	no	no	cellular	7	jul	148	1	152	1	other	no
22	38	managemt	divorced	unknown	no	0	yes	no	cellular	18	nov	96	2	-1	0	unknown	no
23	42	managemt	divorced	tertiary	no	16	no	no	cellular	19	nov	140	3	-1	0	unknown	no
24	44	services	single	secondary	no	106	no	no	unknown	12	jun	109	2	-1	0	unknown	no
25	44	entrepreni	married	secondary	no	93	no	no	cellular	7	jul	125	2	-1	0	unknown	no

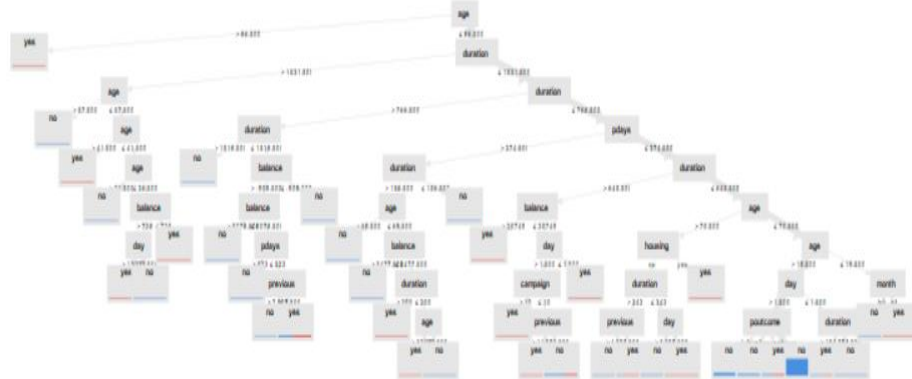
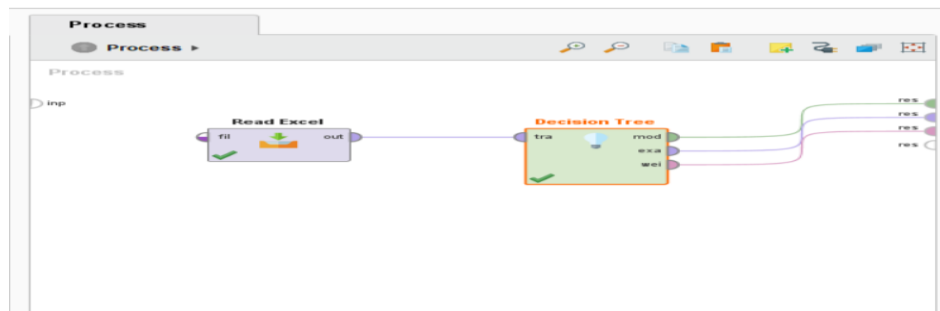
Tabel 4.2 Dataset Penelitian

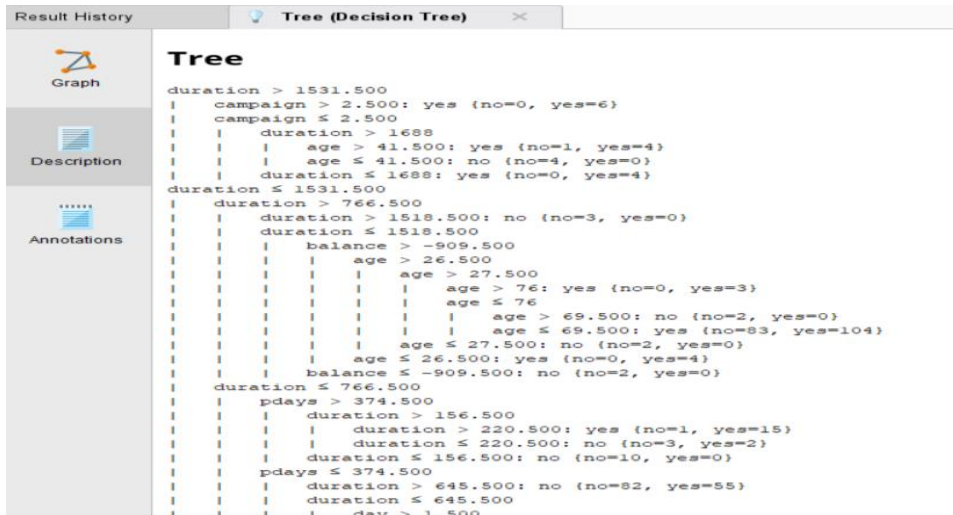
Tabel diatas menunjukkan karakter data pada penelitian ini, data tersebut memiliki label pada atribut Y (deposito) yang berarti data ini merupakan data supervised karena memiliki label, label pada data ini adalah Yes/No deposito. Alat yang digunakan pada penelitian ini adalah aplikasi Rapidminer.

## 4.4 Pemodelan

### 4.4.1 Penelitian Menggunakan Algoritma Decision Tree C4.5

Pada algoritma decision tree, record yang sudah di import ke rapid miner digunakan untuk menentukan pola pohon keputusan. Adapun hasil dari pola pohon keputusan nya sebagai berikut.



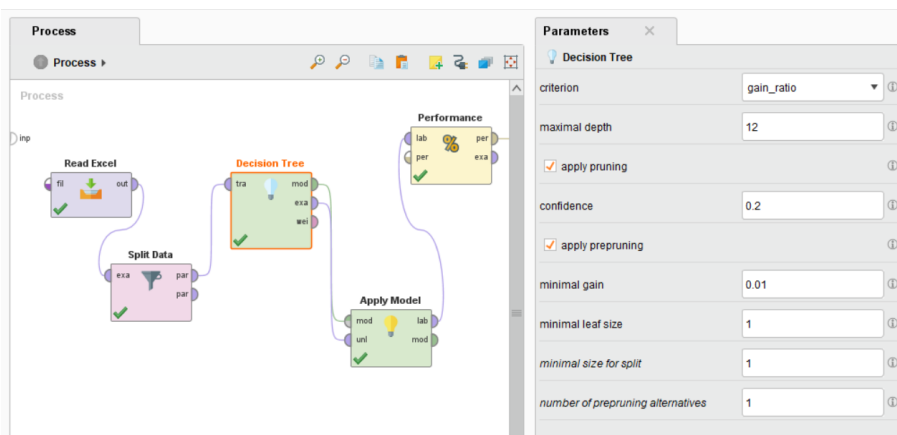


Gambar 4.1. Desain model dan pola pohon keputusan

Setelah diketahui pola pohon keputusan berdasarkan data training, maka akan dimodelkan untuk meningkatkan tingkat akurasi.

#### 4.4.2 Penelitian Menggunakan Algoritma Decision Tree

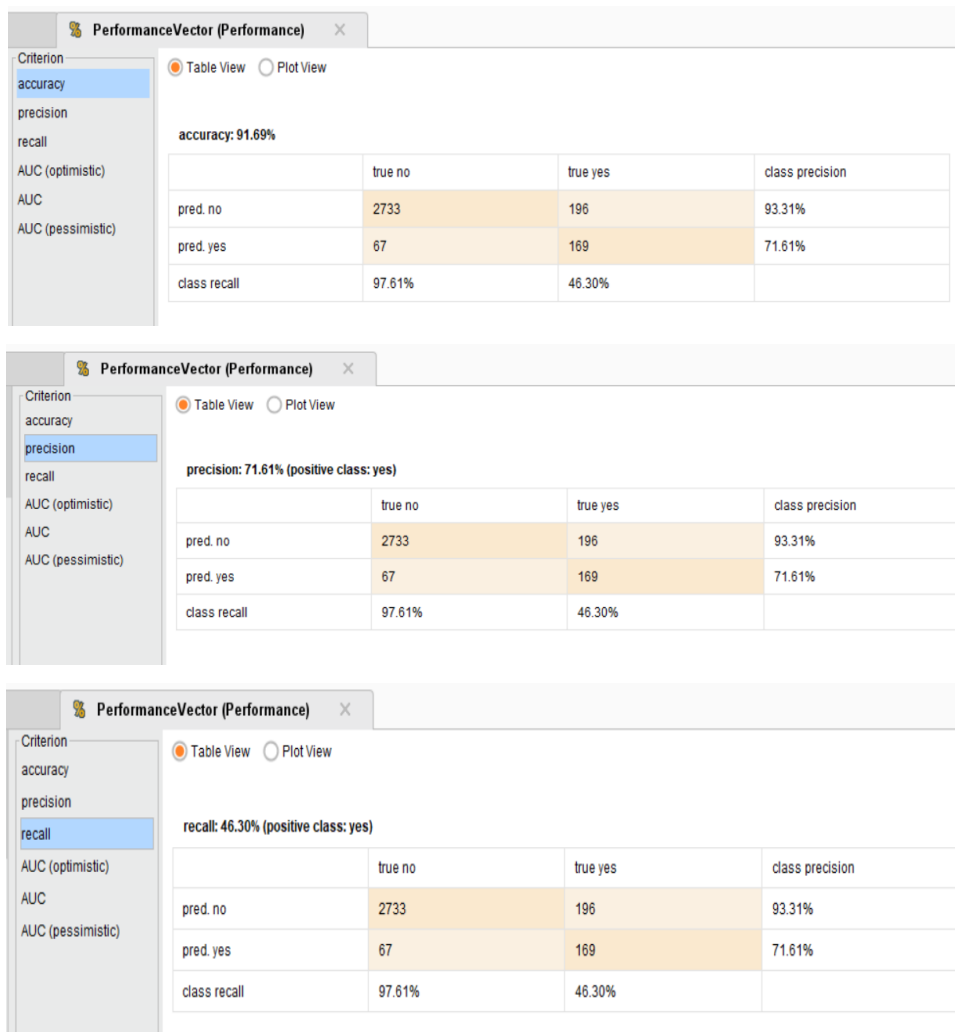
Penerapan data pada Rapidminer untuk klasifikasi calon nasabah deposito bank menggunakan Algoritma Decision Tree ditunjukkan pada gambar dibawah ini:



Gambar 4.2 Penerapan data klasifikasi calon nasabah deposito bank menggunakan Algoritma Decision Tree pada Rapidminer

Pada gambar 4.2 data yang sudah disiapkan diterapkan pada aplikasi rapidminer dan data yang digunakan adalah data sekunder dan algoritma

yang digunakan adalah decision tree. Hasil eksperimen dapat kita lihat pada gambar 4.3 dibawah ini.



Gambar 4.3 Confusion Matrix Hasil klasifikasi calon nasabah deposito bank menggunakan Algoritma Decision Tree pada Rapidminer

Gambar 4.3 adalah confusion matrix yang menunjukkan hasil eksperimen, didalam confusion matrix kita dapat melihat hasil akurasi, class presisi, dan class recall. Akurasi yang dihasilkan adalah 91.69 %, class presisi 71.61 % dan class recall 46.30 %.

Hasil akurasi pada penelitian sebelumnya dengan eksperimen penelitian ini dengan menggunakan algoritma decision tree untuk klasifikasi calon nasabah deposito bank ditunjukkan pada tabel dibawah ini:

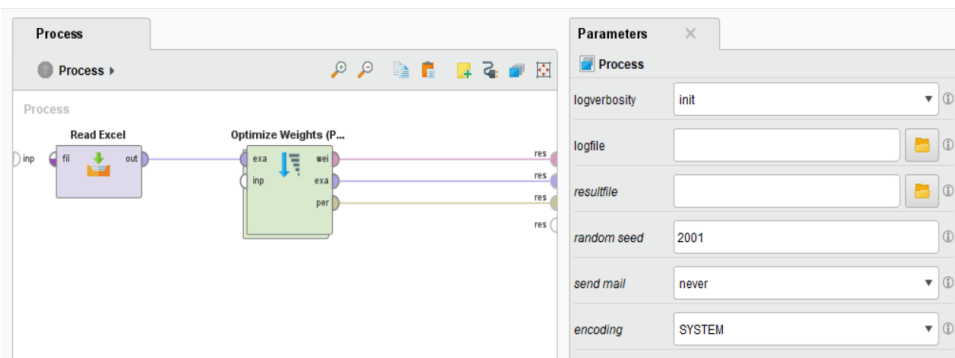
Tabel. 4.3 Perbandingan Hasil akurasi Penelitian sebelumnya dengan penelitian saat ini:

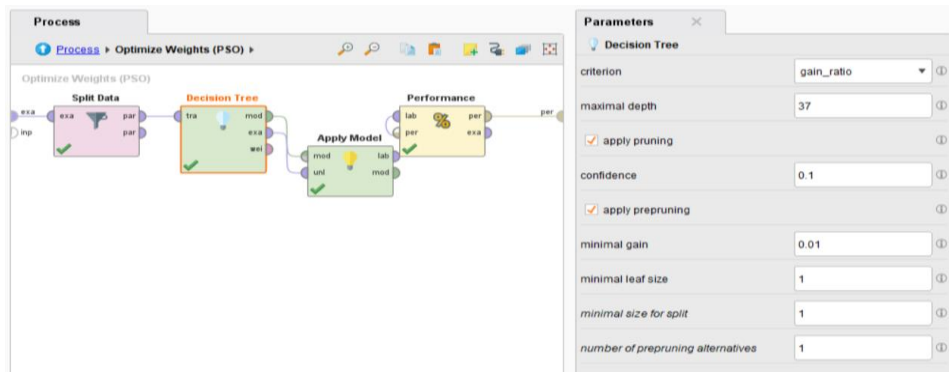
Penelitian	Tingkat Akurasi
Penelitian Sebelumnya (Nuraeni 2021)	<b>91.26%</b>
Penelitian Saat Ini	<b>91.69%</b>

Eksperimen penelitian ini menghasilkan tingkat akurasi tinggi dan tingkat akurasi lebih tinggi dari penelitian sebelumnya, akan tetapi masih dapat ditingkatkan tingkat akurasinya.

#### 4.4.3 Penelitian Menggunakan Particle Swarm Optimization (PSO) dan Decision Tree C4.5

Metode yang dapat meningkatkan tingkat akurasi dalam klasifikasi adalah penggunaan teknik optimasi. Eksperimen lanjutan yang dilakukan pada penelitian ini adalah penggunaan teknik optimasi seleksi fitur menggunakan *Particle Swarm Optimization* (PSO) dan algoritma *Decision Tree C4.5*. Penerapan Metode dan algoritma pada rapidminer ditunjukkan pada gambar di bawah ini:





Gambar 4.4 Penerapan Data Kalsifikasi calon nasabah deposito bank Menggunakan Particle Swarm Optimization (PSO) dan C.45

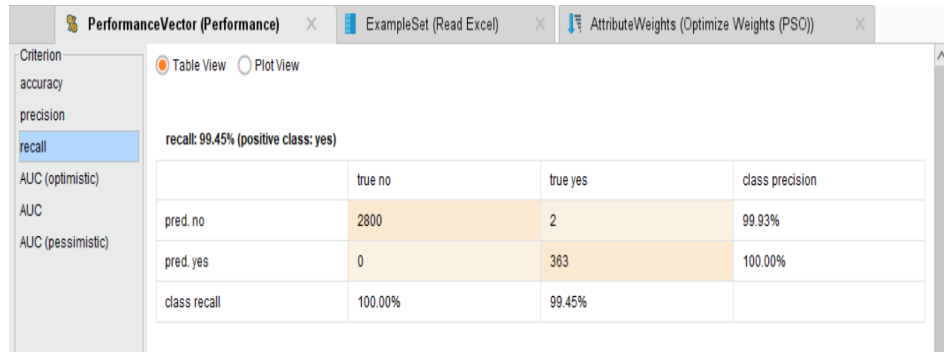
Pada gambar 4.4 data yang sudah disiapkan diterapkan pada aplikasi rapidminer dan penggunaan teknik optimasi seleksi fitur *Particle Swarm Optimization* (PSO) dan algoritma *Decision Tree* C4.5. Hasil eksperimen dapat kita lihat pada gambar 4.5 dibawah ini.

The top screenshot shows the 'PerformanceVector (Performance)' window with the 'accuracy' criterion selected. The overall accuracy is 99.94%. The classification results are as follows:

	true no	true yes	class precision
pred. no	2800	2	99.93%
pred. yes	0	363	100.00%
class recall	100.00%	99.45%	

The bottom screenshot shows the same window with the 'precision' criterion selected. The overall precision is 100.00% (positive class: yes). The classification results are the same as in the top screenshot:

	true no	true yes	class precision
pred. no	2800	2	99.93%
pred. yes	0	363	100.00%
class recall	100.00%	99.45%	



Gambar 4.5. Confusion Matrix Hasil klasifikasi calon nasabah deposito bank menggunakan seleksi fitur PSO dan Algoritma Decision Tree C4.5 pada Rapidminer

Gambar 4.5 adalah confusion matrix yang menunjukkan hasil eksperimen lanjutan, didalam confusion matrix kita dapat melihat hasil akurasi 99.94%, class presisi 100% dan class recall 99.45%.

Tabel 4.4 Perbandingan Jumlah Maximal Depth PSO+Decision Tree

Metode	Prese ntase Kinerj a	Jumlah Maximal Depth							
		5	10	15	20	25	30	35	37
PSO+De cision Tree	Akura si (%)	90.1	92.32	96.4	97.8	99.0	99.8	99.9	99.9
	Presi si (%)	4%	%	0%	5%	2%	1%	1%	4%
	Recall (%)	24.9	46.03	83.5	81.3	91.5	98.3	99.1	99.4
		3%	%	6%	7%	1 %	6 %	8 %	5%

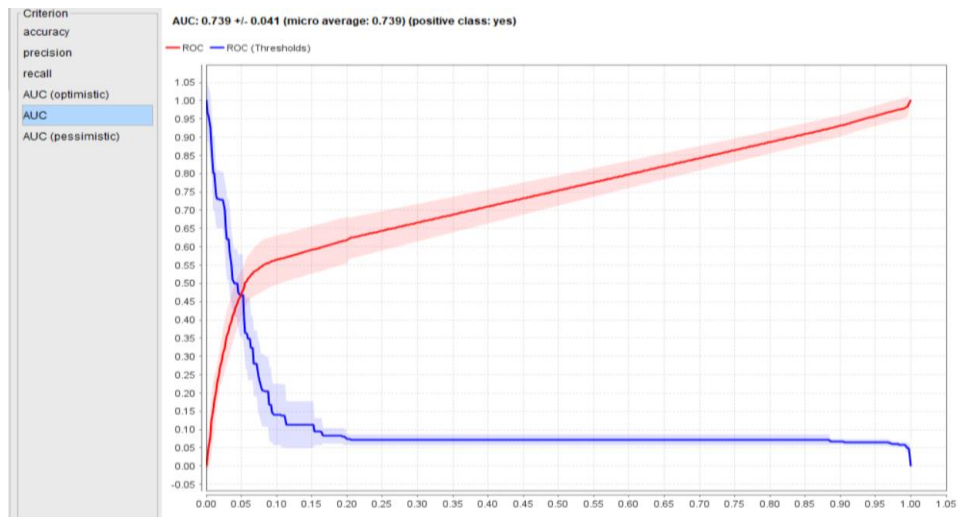
Berdasarkan tabel diatas, menunjukkan bahwa perubahan jumlah Maximal Depth pada algoritma *Decision tree* dalam melakukan klasifikasi berpengaruh dalam menghasilkan akurasi tersebut, akurasi, presisi dan recall yang terbaik dapat kita lihat pada Maximal Depth 37, berbeda dengan hasil akurasi, presisi dan recal lainnya yang tidak lebih tinggi dari hasil Maximal Depth lainnya.

Perbandingan hasil eksperimen yang ditunjukkan pada tabel dibawah ini:

Tabel 4.5 Perbandingan Hasil Akurasi Penggunaan PSO

Penelitian	Tingkat Akurasi
(Agustina 2018)	<b>89.72 %</b>
Penelitian Saat ini	<b>99.94 %</b>

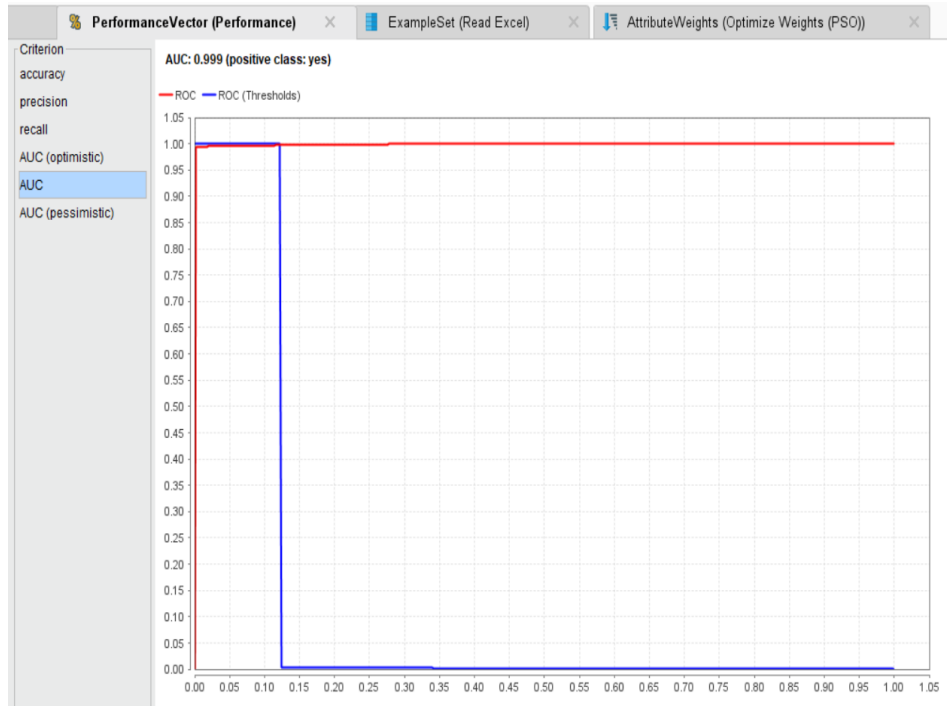
Hasil dari eksperimen ini menghasilkan tingkat akurasi yang tinggi sebesar 99,94%, class presisi 100% dan class recall 98.45% dari penelitian sebelumnya sebesar 89,72%(Agustina 2018). akurasi ini meningkat ketika dilakukannya seleksi fitur oleh algoritma *decision tree* C4.5 pada dataset deposito bank. Selain Confusion matrix untuk mengetahui kinerja dari eksperimen ini kita dapat mengandalkan kurva AUC yang dihasilkan. Perbandingan hasil Kurva AUC pada penelitian sebelumnya dan penelitian saat ini dapat kita lihat pada gambar 4.6 dan 4.7 dibawah ini:



Gambar 4.6 Kurva AUC hasil eksperimen penelitian sebelumnya

Gambar 4.7 menunjukan kinerja algoritma dalam mengklasifikasi calon nasabah deposito bank, kinerja algoritma pada eksperimen ini cukup baik dengan menghasilkan Area Under Curva (AUC) 73%( Fair Classification ).





Gambar 4.7 Kurva AUC hasil eksperimen penelitian saat ini

Gambar 4.7 menunjukkan kinerja algoritma dalam mengklasifikasi calon nasabah deposito bank, kinerja algoritma pada eksperimen ini menghasilkan Area Under Curva (AUC) 99% yang membuktikan bahwa akurasi pada model optimasi ini memiliki tingkat klasifikasi yang sangat baik. (*Excelent Clasification*).

#### 4.4.4 Perbandingan Penelitian Tidak Menggunakan PSO dan Menggunakan PSO

Perbandingan hasil eksperimen yang tidak menggunakan seleksi fitur PSO dan eksperimen yang menggunakan seleksi fitur PSO ditunjukkan pada tabel dibawah ini:

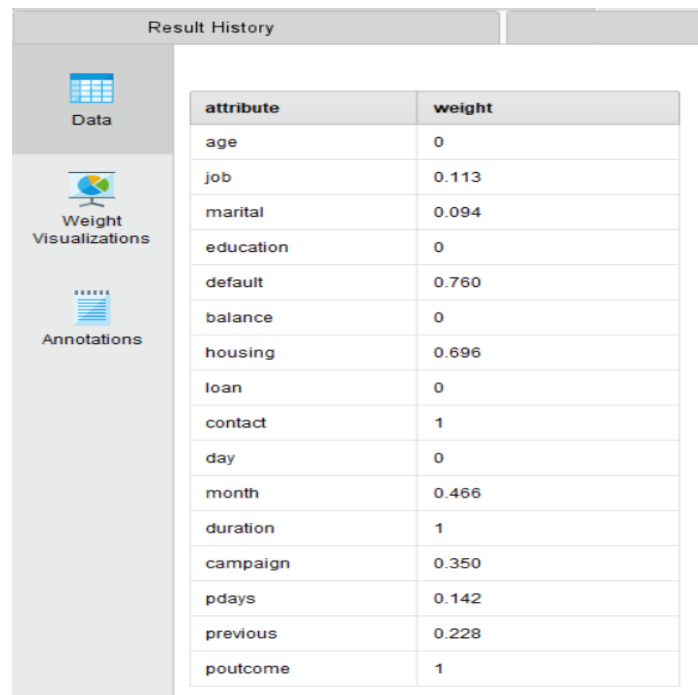
Tabel 4.6 Perbandingan Hasil Akurasi

Penelitian	Tingkat akurasi
Eksperimen Tidak Menggunakan seleksi fitur PSO	<b>99.80 %</b>
Eksperimen Yang Menggunakan seleksi fitur PSO	<b>99.94 %</b>

#### 4.5 Evaluasi

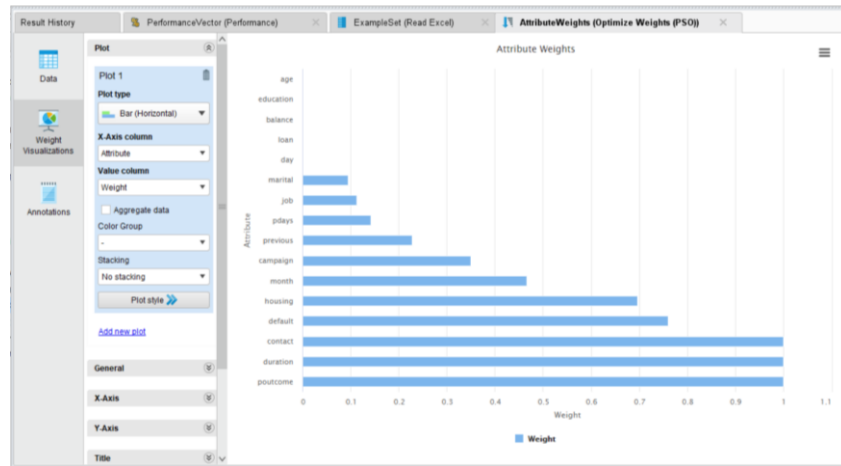
Pemodelan yang telah dilakukan berhasil menghasilkan akurasi yang tinggi. Penggunaan seleksi fitur *Particle Swarm Optimization* (PSO) berhasil meningkatkan akurasi dalam klasifikasi calon nasabah deposito bank.

Berdasarkan model proses yang telah berhasil dilakukan, maka diperoleh data seleksi pembobotan atribut sebagai berikut:



attribute	weight
age	0
job	0.113
marital	0.094
education	0
default	0.760
balance	0
housing	0.696
loan	0
contact	1
day	0
month	0.466
duration	1
campaign	0.350
pdays	0.142
previous	0.228
poutcome	1

Gambar 4.8 Hasil Seleksi Atribut



Gambar 4.9 Hasil Seleksi Atribut

Dari gambar diatas dapat dilihat hasil pembobotan atribut-atribut, terdapat 5 atribut yang memiliki nilai 0 yang artinya atribut tersebut otomatis terseleksi dan tidak perlu digunakan karena tidak memiliki pengaruh dalam proses.

Berikut hasil perbandingan jurnal-jurnal yang telah saya review menggunakan seleksi fitur pada tabel dibawah ini:

Tabel 4.7 Hasil Penelitian

Seleksi Fitur	Metode	Hasil Akurasi
Tanpa Seleksi Fitur (Nuraeni 2021)	Decision Tree	91.26%
Optimize Selection (Evolutionary) (Fauzi et al. 2021)	Naive Bayes	90.18%
GA (Sistem 2021)	Decision Tree	88,30%
Penelitian saat ini (PSO)	Decision Tree	99.84 %

Berdasarkan hasil yang diperlihatkan pada tabel 4.7, diketahui bahwa pemilihan seleksi fitur yang tepat dapat meningkatkan akurasi yang tinggi, seleksi fitur yang digunakan menggunakan *Particle Swarm Optimization* (PSO) dan Algoritma *Decision Tree* C4.5