

BAB 5

HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan empat skenario pengujian pada *dataset glass, lymfografi, vehicle, thyroid* dan *wine* sebagai berikut:

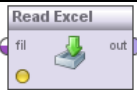
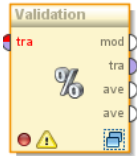
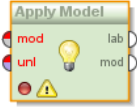
1. Pengujian metode *decision tree* (ID3, C4.5 dan C5.0) menggunakan *dataset* awal.
2. Pengujian metode *decision tree* (ID3, C4.5 dan C5.0) menggunakan *dataset* hasil seleksi dengan algoritma *particle swarm optimazation* (PSO).
3. Pengujian menggunakan metode *cost sensitive decision tree* (ID3, C4.5 dan C5.0) dengan *dataset* hasil seleksi algoritma *particle swarm optimazation* (PSO).
4. Pengujian menggunakan metode *cost sensitive decision tree* C5.0 dan *cost sensitive naïve bayes*

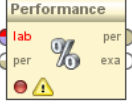


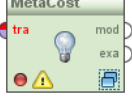


Hasil dari pengujian selanjutnya dilakukan evaluasi untuk mengukur kinerja *classifier* dengan menggunakan parameter *accuracy, recall, precision, f-measure* dan *total cost*.

5.1 Pengenalan Operator *Rapid Minner*

Pada penelitian ini perancangan metode hingga proses pengujian metode menggunakan perangkat lunak *rapid minner 5.3.0*. Perangkat tersebut berfungsi sebagai *machine learning* pada proses penelitian ini. Adapun operator yang digunakan dalam perancangan sistem pengujian antara lain:

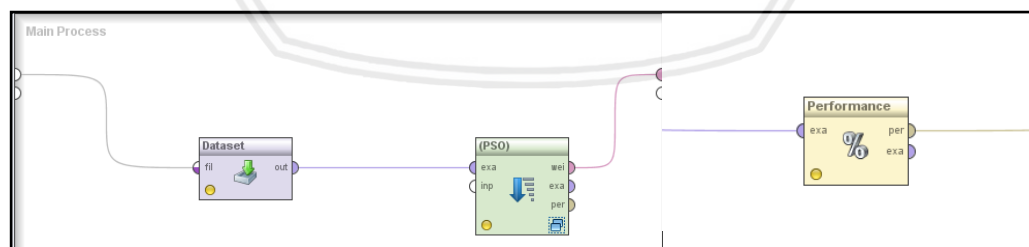
Tabel 5.1 Operator *Rapid Minner*

OPERATOR	FUNGSI
	Operator ini digunakan untuk membaca <i>dataset</i> dengan spesifikasi file excel (.xls)
	Operator ini digunakan untuk membagi <i>dataset</i> menjadi data pelatihan (<i>training</i>) dan pengujian (<i>testing</i>) yang akan digunakan untuk proses pembelajaran dan pengujian.
	Operator ini digunakan untuk mengaplikasikan model pembelajaran dan pelatihan pada <i>dataset</i> .

	<p>Operator ini digunakan untuk mengukur performa dari hasil klasifikasi. Operator ini menampilkan keluaran berupa tabel <i>confusion matrix</i> atau tabel <i>weight</i>.</p>
	<p>Operator ini merupakan aplikasi dari metode <i>decision tree</i> untuk menghasilkan model <i>decision tree</i></p>
	<p>Operator ini merupakan aplikasi dari metode <i>naïve bayes</i> untuk menghasilkan model klasifikasi <i>naïve bayes</i></p>
	<p>Operator ini merupakan model <i>metaclassifier</i> yang menggunakan metode <i>cost sensitive</i> yang menghasilkan <i>cost matrix</i> untuk menghitung label prediksi berdasarkan <i>cost</i> klasifikasi</p>
	<p>Operator <i>adacost</i> adalah algoritma <i>meta</i> yang digunakan untuk meningkatkan performa dengan menghitung <i>cost misclassification</i></p>
	<p>Operator ini menampilkan nilai bobot atribut berdasarkan perhitungan algoritma <i>particle swarm optimization</i></p>

5.2 Proses Seleksi Atribut

Rancangan pengujian menggunakan *rapid minner* untuk menyeleksi atribut ditampilkan seperti gambar 5.1.



Gambar 5.1 Rancangan Proses Seleksi Atribut

Proses awal memasukkan *dataset* kemudian *dataset* diproses menggunakan algoritma *particle swarm optimazation* (PSO) dengan memasukan parameter awal *inertia weight* =1.0 ,

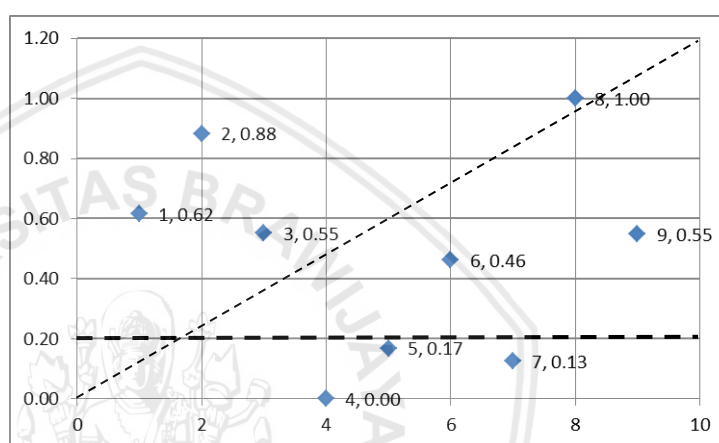
local best weight =1.0, *global best weight* = 0.0, *min weight*= 0.0, *max weight* = 1.0, *population size* = 5, dan *max number of generate* = 10.

5.2.1 Seleksi Atribut *Dataset Glass*

Dataset glass memiliki *atribut* sebanyak sepuluh antara lain *Id number*, RI, Na, Mg, Al, Si, K, Ca, Ba, Fe. Sepuluh atribut tersebut kemudian dilakukan pengujian menggunakan algoritma *particle swarm optimazation* untuk menghitung bobot setiap atribut. Hasil pengujian ditampilkan seperti Tabel 5.2.

Tabel 5.2 Nilai Atribut *Glass*

No	Atribut	Nilai
1	RI	0.62
2	Na	0.88
3	Mg	0.55
4	Al	0.00
5	Si	0.17
6	K	0.46
7	Ca	0.13
8	Ba	1.00
9	Fe	0.55



Gambar 5.2 Posisi berdasarkan nilai atribut *Glass*

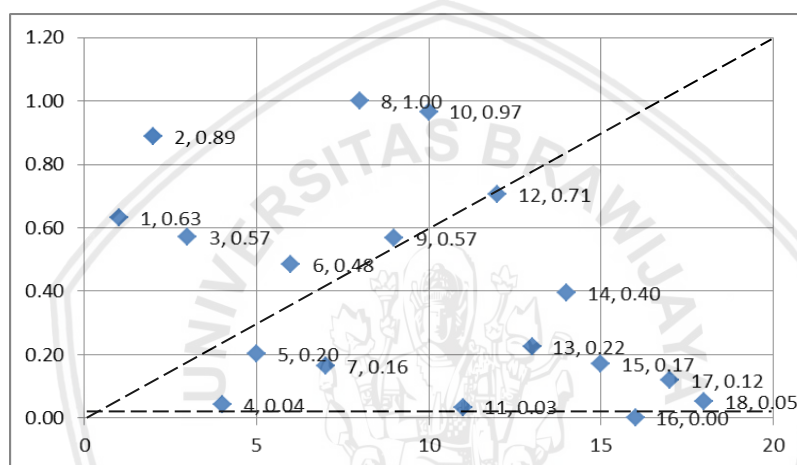
Hasil pengujian selanjutnya dilakukan evaluasi pada atribut yang memiliki nilai bobot 0 atau mendekati 0 sehingga diperoleh atribut yang relevan sehingga mampu meningkatkan performa *classifier* seperti pada Gambar 5.2. Atribut dari *dataset glass* yang dipilih berjumlah enam antara lain RI, Na, Mg, K, Ba, Fe. Keenam atribut tersebut selanjutnya akan digunakan untuk proses penelitian sebagai *dataset glass* PSO.

5.2.2 Seleksi Atribut *Dataset Lympografi*

Dataset lympografi memiliki *atribut* sebanyak 18 antara lain *lymphatics*, *block of affere*, *bl. Of lymph. c,bl. Of lymph. S, by pass, extravasates, regeneration, early uptake in, lym nodes dimin, lym nodes enlar, changes in lym, defect in node, changes in node, changes in stru, special forms, dislocation, exclusion of no, no of nodes in*. Kemudian atribut tersebut dilakukan pengujian menggunakan algoritma *particle swarm optimazation* untuk menghitung bobot setiap atribut. Hasil pengujian ditampilkan pada Tabel 5.3.

Tabel 5.3 Nilai Atribut *Lympografi*

No	Atribut	Nilai	No	Atribut	Nilai
1	Lymphatics	0.63	10	lym nodes enlar	0.97
2	block of affere	0.89	11	changes in lym	0.03
3	bi of lymph	0.57	12	defect in node	0.71
4	bl of lymph	0.04	13	changes in node	0.22
5	by pass	0.20	14	changes in stru	0.40
6	Extravasates	0.48	15	special forms	0.17
7	regeneration of	0.16	16	dislocatio of	0.00
8	early uptake in	1.00	17	exclusion	0.12
9	lym.nodes dimin	0.57	18	no of nodes	0.05

Gambar 5.3 Posisi berdasarkan nilai atribut *Lympografi*

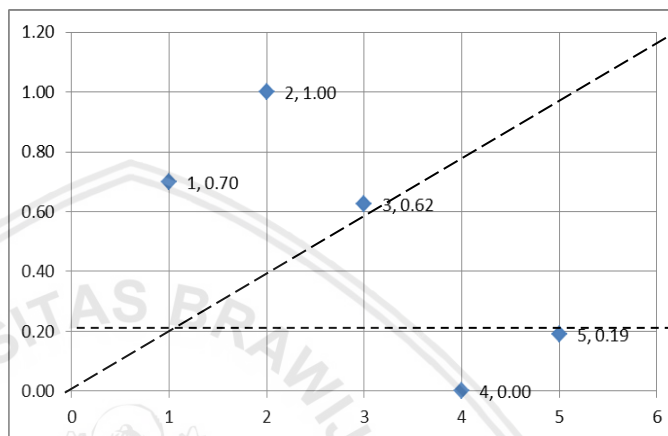
Hasil pengujian selanjutnya dilakukan evaluasi pada atribut yang memiliki nilai bobot 0 atau mendekati 0 sehingga diperoleh atribut yang relevan sehingga mampu meningkatkan performa *classifier*. Atribut dari *dataset lympografi* yang dipilih berjumlah tujuh belas antara lain *lymphatics*, *block of affere*, *bl. Of lymph. c*, *bl. Of lymph. S*, *by pass*, *extravasates*, *regeneration*, *early uptake in*, *lym nodes dimin*, *lym nodes enlar*, *changes in lym*, *defect in node*, *changes in node*, *changes in stru*, *special forms*, *exclusion of no*, *no of nodes in*. Ketujuh belas atribut tersebut selanjutnya akan digunakan untuk proses penelitian sebagai *dataset lympografi PSO*.

5.2.3 Seleksi Atribut *Dataset Thyroid*

Dataset thyroid memiliki atribut sebanyak 5 antara lain *T3-resin uptake test*, *total serum thyroxin*, *total serum triiodothyronine*, *basal thyroid-stimulating hormone (TSH)*, *Maximal absolute difference of TSH*. Kemudian atribut tersebut dilakukan pengujian menggunakan algoritma *particle swarm optimization* untuk menghitung bobot setiap atribut. Hasil pengujian ditampilkan pada Tabel 5.4.

Tabel 5.4 Nilai Atribut *Thyroid*

No.	Atribut	Nilai
1	T3-resin uptake test	0.70
2	Serum thyroxin	1.00
3	serum triiodothyronine	0.19
4	thyroid-stimulating hormone (TSH)	0.00
5	difference of TSH value	0.62



Gambar 5.4 Posisi berdasarkan nilai atribut *Thyroid*

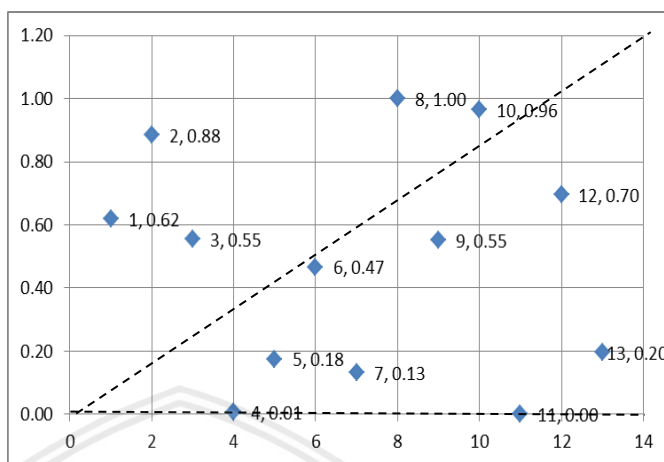
Hasil pengujian selanjutnya dilakukan evaluasi pada atribut yang memiliki nilai bobot 0 atau mendekati 0 sehingga diperoleh atribut yang relevan sehingga mampu meningkatkan performa *classifier*. Atribut dari *dataset thyroid* yang dipilih berjumlah empat antara lain *T3-resin uptake test*, *total serum thyroxin*, *total serum triiodothyronine*, *Maximal absolute difference of TSH*. Keempat atribut tersebut selanjutnya akan digunakan untuk proses penelitian sebagai *dataset thyroid*.

5.2.4 Seleksi Atribut *Dataset Wine*

Dataset wine memiliki atribut sebanyak 13 antara lain *alcohol*, *malic acid*, *ash*, *alkalinity of ash*, *magnesium*, *total phenols*, *flavonoids*, *nonflavanoid phenols*, *proanthocyanins*, *color intensity*, *hue*, *OD280 of diluter wines*, *proline*. Kemudian atribut tersebut dilakukan pengujian menggunakan algoritma *particle swarm optimization* untuk menghitung bobot setiap atribut. Hasil pengujian ditampilkan pada Tabel 5.5.

Tabel 5.5 Nilai Atribut *Wine*

No	Atribut	Nilai
1	alcohol	0.62
2	Malic acid	0.88
3	ash	0.55
4	alkalinity of ash	0.01
5	magnesium	0.18
6	total phenols	0.47
7	flavanoids	0.13
8	nonflavanoid phenols	1.00
9	proanthocyanins	0.55
10	color intensity	0.96
11	hue	0.00
12	diluted wines	0.70
13	proline	0.20

**Gambar 5.5** Posisi berdasarkan nilai atribut *Wine*

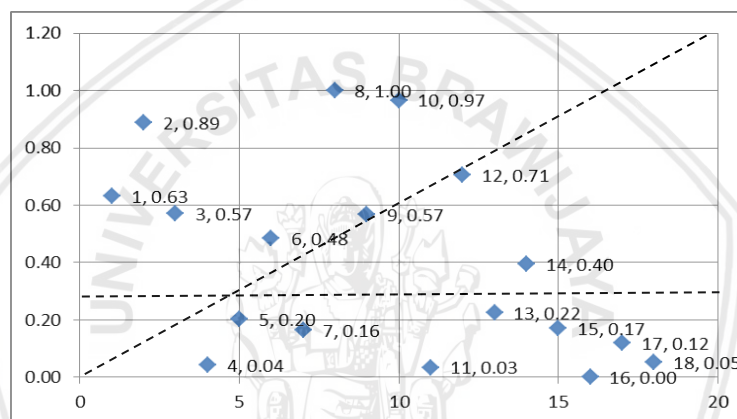
Hasil pengujian selanjutnya dilakukan evaluasi pada atribut yang memiliki nilai bobot 0 atau mendekati 0 sehingga diperoleh atribut yang relevan sehingga mampu meningkatkan performa *classifier*. Atribut dari *dataset wine* yang dipilih berjumlah dua belas antara lain *alcohol*, *malic acid*, *ash*, *alkalinity of ash*, *magnesium*, *total phenols*, *flavonoids*, *nonflavanoid phenols*, *proanthocyanins*, *color intensity*, *OD280 of diluter wines*, *proline*. Ke dua belas atribut tersebut selanjutnya akan digunakan untuk proses penelitian sebagai *dataset wine PSO*.

5.2.5 Seleksi Atribut *Dataset Vehicle*

Dataset vehicle memiliki atribut sebanyak 18 antara lain *compactness*, *circularity*, *distance circularity*, *radius ratio*, *pr. Axis aspect ratio*, *max length aspect ratio*, *scatter ratio*, *elongatedness*, *pr axis rectangularity*, *max length rectangularity*, *scaled variance*, *scaled variance*, *scaled radius of gyration*, *skewness about*, *skewness about*, *kurtosis about*, *kurtosis about*, *hollows ratio*. Kemudian atribut tersebut dilakukan pengujian menggunakan algoritma *particle swarm optimazation* untuk menghitung bobot setiap atribut. Hasil pengujian ditampilkan pada Tabel 5.6.

Tabel 5.6 Nilai Atribut *Vehicle*

No	Atribut	Nilai	No	Atribut	Nilai
1	COMPACTNESS	0.63	10	MAX.LENGTH RECTANGULARITY	0.97
2	CIRCULARITY	0.89	11	SCALED VARIANCE	0.03
3	DISTANCE CIRCULARITY	0.57	12	SCALED VARIANCE_1	0.71
4	RADIUS RATIO	0.04	13	SCALED RADIUS OF GYRATION	0.22
5	PR.AXIS ASPECT RATIO	0.20	14	SKEWNESS ABOUT	0.40
6	MAX.LENGTH ASPECT RATIO	0.48	15	SKEWNESS ABOUT_1	0.17
7	SCATTER RATIO	0.16	16	KURTOSIS ABOUT	0.00
8	ELONGATEDNESS	1.00	17	KURTOSIS ABOUT_1	0.12
9	PR.AXIS RECTANGULARITY	0.57	18	HOLLOWS RATIO	0.05

Gambar 5.6 Posisi berdasarkan nilai atribut *Vehicle*

Hasil pengujian selanjutnya dilakukan evaluasi pada atribut yang memiliki nilai bobot 0 atau mendekati 0 sehingga diperoleh atribut yang relevan sehingga mampu meningkatkan performa *classifier*. Atribut dari *dataset vehicle* yang dipilih berjumlah sembilan antara lain *compactness*, *circularity*, *distance circularity*, *pr. max length aspect ratio*, *elongatedness*, *pr axis rectangularity*, *max length rectangularity*, *scaled variance*, *skewness about*. Kesembilan atribut tersebut selanjutnya akan digunakan untuk proses penelitian sebagai *dataset vehicle* PSO.

5.3 Pengujian Penelitian

Pengujian penelitian pada *dataset glass*, *lympografi*, *thyroid*, *vehicle* dan *wine* dilakukan menggunakan empat skenario sebagai berikut:

Skenario 1: Pengujian algoritma ID3, C4.5 dan C5.0 dengan menggunakan *dataset* awal.

Skenario 2: Pengujian algoritma ID3, C4.5 dan C5.0 dengan menggunakan *dataset* hasil seleksi menggunakan PSO.

Skenario 3: Pengujian dengan menggunakan metode *cost sensitive decision tree* dengan mengimplementasikan algoritma ID3, C4.5 dan C5.0.

Skenario 4: Pengujian dengan membandingkan *cost sensitive decision tree C5.0* dengan *cost sensitive naïve bayes*.

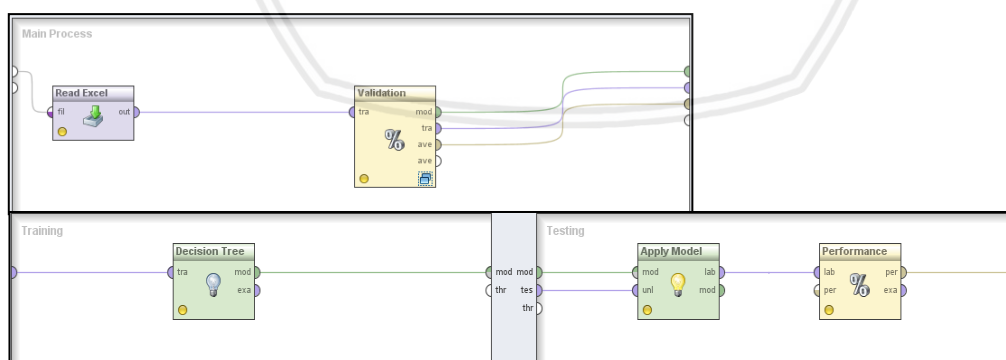
Keempat skenario tersebut kemudian dirancang menggunakan *rapid minner 5.3.0* dengan menggunakan metode *10 Fold Cross Validation* yaitu membagi *dataset* menjadi 90% data pelatihan dan 10% data pengujian. Berikut merupakan rancangan pengujian metode menggunakan *rapid minner*.

5.4 Perancangan Skenario Pengujian

Pada penelitian ini menggunakan 4 skenario pengujian yang telah dijelaskan pada bab 5.3. Keempat skenario tersebut dirancang menggunakan *rapid miner 5.3.0*. Model perancangan pengujian digambarkan sebagai berikut.

5.4.1 Perancangan Skenario 1

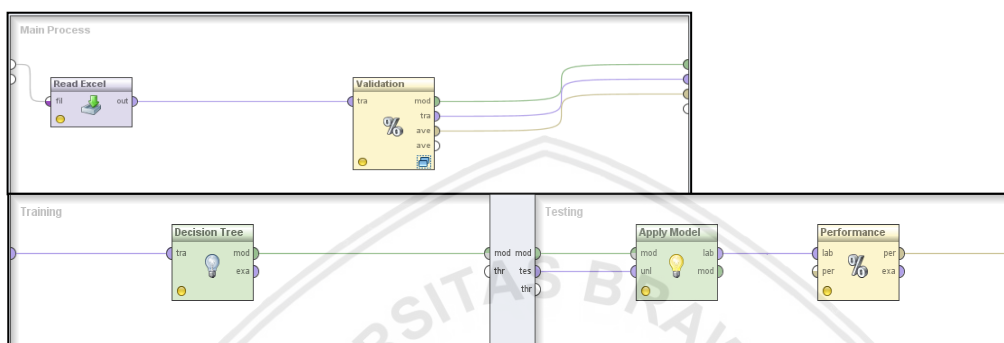
Pada perancangan skenario 1 *dataset* pengujian menggunakan *dataset* tanpa melalui *preprocessing*. Selanjutnya *dataset* tersebut dilakukan *validation* dengan menggunakan metode *10 cross validation* yaitu membagi 90% data pelatihan dan 10% data pengujian. Kemudian *dataset* dilakukan pembelajaran menggunakan metode *decision tree* (ID3, C4.5 dan C5.0). Hasil akhir akan ditampilkan dalam bentuk tabel *confusion matrix*. Rancangan sistem ditampilkan pada gambar 5.7.



Gambar 5.7 Perancangan Pengujian Skenario 1

5.4.2 Perancangan Skenario 2

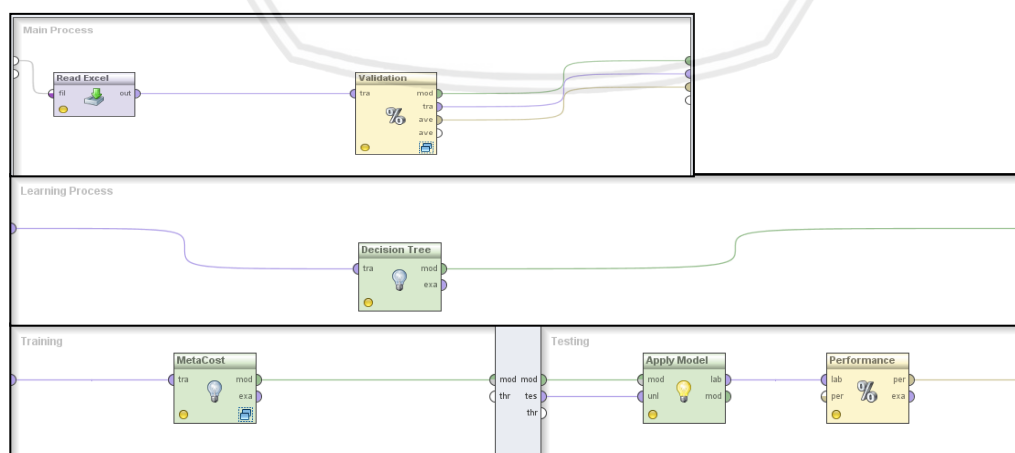
Pada perancangan skenario 2 *dataset* pengujian menggunakan *dataset* PSO. Selanjutnya *dataset* tersebut dilakukan *validation* dengan menggunakan metode 10 *cross validation* yaitu membagi 90% data pelatihan dan 10% data pengujian. Kemudian *dataset* dilakukan pembelajaran menggunakan metode *decision tree* (ID3, C4.5 dan C5.0). Hasil akhir akan ditampilkan dalam bentuk tabel *confusion matrix*. Rancangan system ditampilkan pada Gambar 5.8.



Gambar 5.8 Perancangan Pengujian Skenario 2

5.4.3 Perancangan Skenario 3

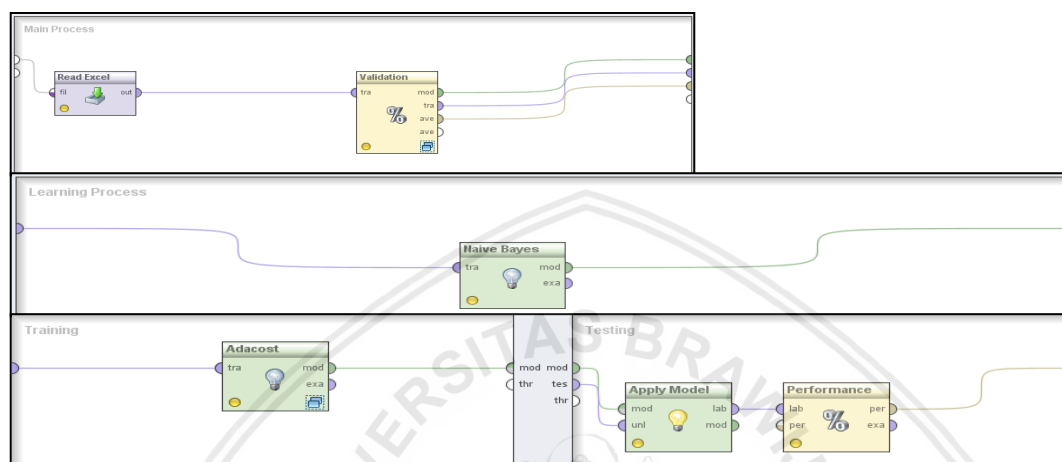
Pada perancangan skenario 3 *dataset* pengujian menggunakan *dataset* PSO. Selanjutnya *dataset* tersebut dilakukan *validation* dengan menggunakan metode 10 *cross validation* yaitu membagi 90% data pelatihan dan 10% data pengujian. Kemudian *dataset* dilakukan pembelajaran menggunakan metode *decision tree* (ID3, C4.5 dan C5.0) dan *metacost*. Hasil akhir akan ditampilkan dalam bentuk tabel *confusion matrix*. Rancangan system ditampilkan pada Gambar 5.9.



Gambar 5.9 Perancangan Pengujian Skenario 3

5.4.4 Perancangan Skenario 4

Pada perancangan skenario 4 *dataset* pengujian menggunakan *dataset* PSO. Selanjutnya *dataset* tersebut dilakukan *validation* dengan menggunakan metode 10 *cross validation* yaitu membagi 90% data pelatihan dan 10% data pengujian. Kemudian *dataset* dilakukan pembelajaran menggunakan metode *naïve bayes* dan *adacost*. Hasil akhir akan ditampilkan dalam bentuk tabel *confusion matrix*. Rancangan sistem ditampilkan pada Gambar 5.10.



Gambar 5.10 Perancangan Pengujian Skenario 4

5.5 Hasil Pengujian *Decision Tree*

Hasil dari perancangan tersebut menghasilkan keluaran berupa tabel *confusion matrix* yang kemudian dapat mengukur *performa classifier* berdasarkan nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Masing-masing tiap pengujian *dataset* sebagai berikut.

5.5.1 Pengujian Algoritma ID3 pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.7.

Tabel 5.7 Hasil pengujian *dataset glass* dengan ID3

	Actual Y1	Actual Y2	Actual Y3	Actual Y5	Actual Y6	Actual Y7
Pred. Y1	70	76	17	13	9	29
Pred. Y2	0	0	0	0	0	0
Pred. Y3	0	0	0	0	0	0
Pred. Y5	0	0	0	0	0	0
Pred. Y6	0	0	0	0	0	0
Pred. Y7	0	0	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 70$, $TN = 926$, $FP = 144$, $FN = 144$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* dengan menggunakan metode *decision tree* ID3 diperoleh nilai *accuracy* sebesar 32.71%, nilai *recall* sebesar 32.71%, nilai *precision* sebesar 86.54%, nilai *f-measure* sebesar 47.48% dan total *cost* sebesar 41472.

5.5.2 Pengujian Algoritma C4.5 pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.8.

Tabel 5.8 Hasil pengujian *dataset glass* dengan C4.5

	Actual Y1	Actual Y2	Actual Y3	Actual Y5	Actual Y6	Actual Y7
Pred. Y1	52	14	6	0	0	1
Pred. Y2	12	47	5	4	1	1
Pred. Y3	4	10	6	0	0	0
Pred. Y5	0	2	0	8	0	1
Pred. Y6	0	1	0	0	8	1
Pred. Y7	2	2	0	1	0	25

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 146$, $TN = 1002$, $FP = 68$, $FN = 68$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* dengan menggunakan metode *decision tree* C4.5 diperoleh nilai *accuracy* sebesar 68.22%,

nilai *recall* sebesar 68.22%, nilai *precision* sebesar 93.64%, nilai *f-measure* sebesar 78.94% dan total cost sebesar 9248.

5.5.3 Pengujian Algoritma C5.0 pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.9.

Tabel 5.9 Hasil pengujian *dataset glass* dengan C5.0

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	49	12	5	0	0	1
pred. Y2	17	49	5	4	1	3
pred. Y3	4	9	7	0	0	1
pred. Y5	0	2	0	8	0	1
pred. Y6	0	1	0	0	8	0
pred. Y7	0	3	0	1	0	23

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 168, TN = 346, FP = 10, FN = 10 sehingga dapat digunakan untuk mencari nilai *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 94.38%, nilai *recall* sebesar 94.38%, nilai *precision* sebesar 97.19%, nilai *f-measure* sebesar 95.77% dan total *cost* sebesar 200.

5.5.4 Pengujian Algoritma ID3 dan PSO pada *Dataset Glass*

Dataset glass PSO yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.10.

Tabel 5.10 Hasil pengujian *dataset glass* dengan ID3 dan PSO

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	70	76	17	13	9	29
pred. Y2	0	0	0	0	0	0
pred. Y3	0	0	0	0	0	0
pred. Y5	0	0	0	0	0	0
pred. Y6	0	0	0	0	0	0
pred. Y7	0	0	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 70$, $TN = 926$, $FP = 144$, $FN = 144$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *decision tree* ID3 diperoleh nilai *accuracy* sebesar 32.71%, nilai *recall* sebesar 32.71%, nilai *precision* sebesar 86.54%, nilai *f-measure* sebesar 47.48% dan *total cost* sebesar 41472.

5.5.5 Pengujian Algoritma C4.5 dan PSO pada *Dataset Glass*

Dataset glass PSO yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.11.

Tabel 5.11 Hasil pengujian *dataset glass* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	54	13	7	0	0	0
pred. Y2	8	48	7	1	1	2
pred. Y3	6	10	3	0	0	0
pred. Y5	0	3	0	10	0	3
pred. Y6	0	1	0	0	8	1
pred. Y7	2	1	0	2	0	23

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 146$, $TN = 1002$, $FP = 68$, $FN = 68$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *decision tree* C4.5 diperoleh nilai *accuracy* sebesar 68.22%,

nilai *recall* sebesar 68.22%, nilai *precision* sebesar 93.64%, nilai *f-measure* sebesar 78.94% dan total *cost* sebesar 9248.

5.5.6 Pengujian Algoritma C5.0 dan PSO pada *Dataset Glass*

Dataset glass PSO yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.12.

Tabel 5.12 Hasil pengujian *dataset glass* dengan C5.0 dan PSO

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	55	14	5	0	0	0
pred. Y2	7	48	5	1	1	3
pred. Y3	8	9	7	0	0	1
pred. Y5	0	2	0	10	0	3
pred. Y6	0	1	0	0	8	0
pred. Y7	0	2	0	2	0	22

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 150, TN = 1006, FP = 64, FN = 64 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 70.09%, nilai *recall* sebesar 70.09%, nilai *precision* sebesar 94.02%, nilai *f-measure* sebesar 80.31% dan total *cost* sebesar 8192.

5.5.7 Pengujian *Cost Sensitive Decision Tree ID3* pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) ID3 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.13.

Tabel 5.13 Hasil pengujian *dataset glass* dengan csDT ID3

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	70	76	17	13	9	29
pred. Y2	0	0	0	0	0	0
pred. Y3	0	0	0	0	0	0
pred. Y5	0	0	0	0	0	0
pred. Y6	0	0	0	0	0	0
pred. Y7	0	0	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 70, TN = 926, FP = 144, FN = 144 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *cost sensitive decision tree* ID3 diperoleh nilai *accuracy* sebesar 32.71%, nilai *recall* sebesar 32.71%, nilai *precision* sebesar 86.54%, nilai *f-measure* sebesar 47.48% dan total *cost* sebesar 41472.

5.5.8 Pengujian *Cost Sensitive Decision Tree C4.5* pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C4.5 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.14.

Tabel 5.14 Hasil pengujian *dataset glass* dengan csDT C4.5

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	53	13	8	0	0	2
pred. Y2	11	51	7	1	1	1
pred. Y3	3	6	2	0	0	0
pred. Y5	0	2	0	11	0	1
pred. Y6	0	2	0	0	8	0
pred. Y7	3	2	0	1	0	25

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 150, TN = 1006, FP = 68, FN = 68 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *cost sensitive decision tree* C4.5 diperoleh nilai *accuracy* sebesar 70.09% ,nilai *recall* sebesar 70.09%, nilai *precision* sebesar 94.02%, nilai *f-measure* sebesar 80.31% dan total *cost* sebesar 8192.

5.5.9 Pengujian *Cost Sensitive Decision Tree C5.0* pada *Dataset Glass*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C5.0 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.15.

Tabel 5.15 Hasil pengujian *dataset glass* dengan csDT C5.0

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	64	10	9	0	0	1
pred. Y2	5	55	7	1	1	1
pred. Y3	1	5	1	0	0	1
pred. Y5	0	2	0	11	0	1
pred. Y6	0	3	0	0	8	1
pred. Y7	0	1	0	1	0	24

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 163$, $TN = 1019$, $FP = 51$, $FN = 51$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) C5.0 diperoleh nilai *accuracy* sebesar 76.17%, nilai *recall* sebesar 76.17%, nilai *precision* sebesar 95.23%, nilai *f-measure* sebesar 84.64% dan total *cost* sebesar 5202.

5.5.10 Pengujian Algoritma ID3 pada *Dataset Lymphografi*

Dataset lymphografi yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.16.

Tabel 5.16 Hasil pengujian *dataset lymphografi* dengan ID3

	true Y1	true Y2	true Y3	true Y5
pred. Y1	61	81	4	1
pred. Y2	0	0	0	1
pred. Y3	0	0	0	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 61$, $TN = 357$, $FP = 87$, $FN = 87$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 41.22%, nilai *recall* sebesar 41.22%, nilai *precision* sebesar 80.41%, nilai *f-measure* sebesar 54.50% dan *total cost* sebesar 15138.

5.5.11 Pengujian Algoritma C4.5 pada *Dataset Lypografi*

Dataset lypografi yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.17.

Tabel 5.17 Hasil pengujian *dataset lypografi* dengan C4.5

	true Y1	true Y2	true Y3	true Y5
pred. Y1	46	16	2	1
pred. Y2	15	65	1	1
pred. Y3	0	0	1	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 112$, $TN = 408$, $FP = 36$, $FN = 36$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 75.68%, nilai *recall* sebesar 75.58%, nilai *precision* sebesar 91.89%, nilai *f-measure* sebesar 83.00% dan *total cost* sebesar 2592.

5.5.12 Pengujian Algoritma C5.0 pada *Dataset Lypografi*

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.18.

Tabel 5.18 Hasil pengujian *dataset lympografi* dengan C5.0

	true Y1	true Y2	true Y3	true Y5
pred. Y1	44	16	2	1
pred. Y2	15	65	0	1
pred. Y3	2	0	2	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 111$, $TN = 407$, $FP = 37$, $FN = 37$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lympografi* dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 75.00%, nilai *recall* sebesar 75.00%, nilai *precision* sebesar 91.67%, nilai *f-measure* sebesar 82.50% dan *total cost* sebesar 2738.

5.5.13 Pengujian Algoritma ID3 dan PSO pada *Dataset Lympografi*

Dataset lympografi PSO yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.19.

Tabel 5.19 Hasil pengujian *dataset lympografi* dengan ID3 dan PSO

	true Y1	true Y2	true Y3	true Y5
pred. Y1	61	81	4	1
pred. Y2	0	0	0	1
pred. Y3	0	0	0	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 61$, $TN = 357$, $FP = 87$, $FN = 87$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lympografi* PSO dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 41.22%, nilai *recall* sebesar 41.22%, nilai *precision* sebesar 80.41%, nilai *f-measure* sebesar 54.50% dan *total cost* sebesar 15138.

5.5.14 Pengujian Algoritma C4.5 dan PSO pada *Dataset Lymphografi*

Dataset lymphografi PSO yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 17 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.20.

Tabel 5.20 Hasil pengujian *dataset lymphografi* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3	true Y5
pred. Y1	43	13	0	1
pred. Y2	14	65	0	0
pred. Y3	0	0	2	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 110$, $TN = 386$, $FP = 28$, $FN = 28$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lymphografi* PSO dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 79.71%, nilai *recall* sebesar 79.71%, nilai *precision* sebesar 93.24%, nilai *f-measure* sebesar 85.94% dan *total cost* sebesar 1568.

5.5.15 Pengujian Algoritma C5.0 dan PSO pada *Dataset Lymphografi*

Dataset lymphografi yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 17 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.21.

Tabel 5.21 Hasil pengujian *dataset lymphografi* dengan C5.0 dan PSO

	true Y1	true Y2	true Y3	true Y5
pred. Y1	40	11	0	0
pred. Y2	17	67	0	1
pred. Y3	0	0	2	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 109$, $TN = 385$, $FP = 29$, $FN = 29$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-*

measure dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 78.99%, nilai *recall* sebesar 78.99%, nilai *precision* sebesar 93.00%, nilai *f-measure* sebesar 85.42% dan *total cost* sebesar 1682.

5.5.16 Pengujian *Cost Sensitive Decision Tree* ID3 pada *Dataset Lypografi*

Dataset lypografi PSO yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 17 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) ID3 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.22.

Tabel 5.22 Hasil pengujian *dataset lypografi* dengan csDT ID3

	true Y1	true Y2	true Y3	true Y5
pred. Y1	61	81	4	1
pred. Y2	0	0	0	1
pred. Y3	0	0	0	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 61, TN = 357, FP = 87, FN = 87 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* PSO dengan menggunakan metode *cost sensitive decision tree* ID3 diperoleh nilai *accuracy* sebesar 41.22%, nilai *recall* sebesar 41.22%, nilai *precision* sebesar 80.41%, nilai *f-measure* sebesar 54.50% dan *total cost* sebesar 15138.

5.5.17 Pengujian *Cost Sensitive Decision Tree* C4.5 pada *Dataset Lypografi*

Dataset lypografi PSO yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 17 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C4.5 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.23.

Tabel 5.23 Hasil pengujian *dataset lypografi* dengan csDT C4.5

	true Y1	true Y2	true Y3	true Y5
pred. Y1	47	12	0	0
pred. Y2	10	66	0	1
pred. Y3	0	0	2	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 115, TN = 391, FP = 23, FN = 23 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* PSO dengan menggunakan metode *cost sensitive decision tree* C4.5 diperoleh nilai *accuracy* sebesar 83.33%, nilai *recall* sebesar 83.33%, nilai *precision* sebesar 94.44%, nilai *f-measure* sebesar 88.54% dan total *cost* sebesar 1058.

5.5.18 Pengujian *Cost Sensitive Decision Tree* C5.0 pada *Dataset Lypografi*

Dataset lypografi PSO yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 17 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C5.0 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.24.

Tabel 5.24 Hasil pengujian *dataset lypografi* dengan csDT C5.0

	true Y1	true Y2	true Y3	true Y5
pred. Y1	45	10	0	0
pred. Y2	12	68	0	1
pred. Y3	0	0	2	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 115, TN = 391, FP = 23, FN = 23 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 83.33%, nilai *recall* sebesar 83.33%, nilai *precision* sebesar 94.44%, nilai *f-measure* sebesar 88.54% dan total *cost* sebesar 1058.

5.5.19 Pengujian Algoritma ID3 pada *Dataset Thyroid*

Dataset thyroid yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 5 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.25.

Tabel 5.25 Hasil pengujian *dataset thyroid* dengan ID3

	true Y1	true Y2	true Y3
pred. Y1	150	35	30
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 150 , TN = 365 , FP = 65, FN = 65 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* dengan menggunakan metode *decision tree* ID3 diperoleh nilai *accuracy* sebesar 69.77%, nilai *recall* sebesar 69.77%, nilai *precision* sebesar 84.88%, nilai *f-measure* sebesar 76.59% dan *total cost* sebesar 8450.

5.5.20 Pengujian Algoritma C4.5 pada *Dataset Thyroid*

Dataset thyroid yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 5 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.26.

Tabel 5.26 Hasil pengujian *dataset thyroid* dengan C4.5

	true Y1	true Y2	true Y3
pred. Y1	143	4	2
pred. Y2	4	31	0
pred. Y3	3	0	28

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 202, TN = 417, FP = 13, FN = 13 sehingga dapat digunakan untuk mencari nilai *accuracy*, *recall*, *precision*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* dengan menggunakan metode *decision tree* C4.5 diperoleh nilai *accuracy* sebesar 93.95%, nilai *recall* sebesar

93.95%, nilai *precision* sebesar 96.98%, nilai *f-measure* sebesar 95.44% dan total *cost* sebesar 338.

5.5.21 Pengujian Algoritma C5.0 pada Dataset Thyroid

Dataset thyroid yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 5 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.27.

Tabel 5.27 Hasil pengujian *dataset glass* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	146	5	3
pred. Y2	1	30	0
pred. Y3	3	0	27

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 203, TN = 418, FP = 12, FN = 12 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 94.42%, nilai *recall* sebesar 94.42%, nilai *precision* sebesar 97.21%, nilai *f-measure* sebesar 95.79% dan total *cost* sebesar 288.

5.5.22 Pengujian Algoritma ID3 dan PSO pada Dataset Thyroid

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 5 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.28.

Tabel 5.28 Hasil pengujian *dataset thyroid* dengan ID3 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	150	35	30
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 150, TN = 365, FP = 65, FN = 65 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-*

measure dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 69.77%, nilai *recall* sebesar 69.77%, nilai *precision* sebesar 84.88%, nilai *f-measure* sebesar 76.59% dan total *cost* sebesar 8450.

5.5.23 Pengujian Algoritma C4.5 dan PSO pada *Dataset Thyroid*

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 5 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.29.

Tabel 5.29 Hasil pengujian *dataset thyroid* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	144	4	3
pred. Y2	3	31	0
pred. Y3	3	0	27

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 202, TN = 417, FP = 13, FN = 13 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 93.95%, nilai *recall* sebesar 93.95%, nilai *precision* sebesar 96.98%, nilai *f-measure* sebesar 95.44% dan total *cost* sebesar 338.

5.5.24 Pengujian Algoritma C5.0 dan PSO pada *Dataset Thyroid*

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.30.

Tabel 5.30 Hasil pengujian *dataset thyroid* dengan C5.0 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	146	5	3
pred. Y2	1	30	0
pred. Y3	3	0	27

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 203$, $TN = 418$, $FP = 12$, $FN = 12$ sehingga dapat digunakan untuk mencari nilai *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 94.42%, nilai *recall* sebesar 94.42%, nilai *precision* sebesar 97.21%, nilai *f-measure* sebesar 95.79% dan total *cost* sebesar 288.

5.5.25 Pengujian Cost Sensitive Decision Tree ID3 pada Dataset Thyroid

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) ID3 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.31.

Tabel 5.31 Hasil pengujian *dataset thyroid* dengan csDT ID3

	true Y1	true Y2	true Y3
pred. Y1	150	35	30
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 150$, $TN = 365$, $FP = 65$, $FN = 65$ sehingga dapat digunakan untuk mencari nilai *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode *cost sensitive decision tree* ID3 diperoleh nilai *accuracy* sebesar 69.77%, nilai *recall* sebesar 69.77%, nilai *precision* sebesar 84.88%, nilai *f-measure* sebesar 76.59% dan total *cost* sebesar 8450.

5.5.26 Pengujian Cost Sensitive Decision Tree C4.5 pada Dataset Thyroid

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C4.5 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.32.

Tabel 5.32 Hasil pengujian *dataset thyroid* dengan csDT C4.5

	true Y1	true Y2	true Y3
pred. Y1	45	10	0
pred. Y2	12	68	0
pred. Y3	0	0	2

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 201, TN = 416, FP = 13, FN = 13 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 93.49%, nilai *recall* sebesar 93.49%, nilai *precision* sebesar 96.74%, nilai *f-measure* sebesar 95.09% dan total *cost* sebesar 392.

5.5.27 Pengujian Cost Sensitive Decision Tree C5.0 pada Dataset Thyroid

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) C5.0 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.33.

Tabel 5.33 Hasil pengujian *dataset thyroid* dengan csDT C5.0

	true Y1	true Y2	true Y3
pred. Y1	147	4	2
pred. Y2	0	31	0
pred. Y3	3	0	28

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 206, TN = 421, FP = 9, FN = 9 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode *cost sensitive decision tree* C5.0 diperoleh nilai *accuracy* sebesar 95.81% ,nilai *recall* sebesar 95.81%, nilai *precision* sebesar 97.91%, nilai *f-measure* sebesar 96.85% dan total *cost* sebesar 162.

5.5.28 Pengujian Algoritma ID3 pada *Dataset Wine*

Dataset wine yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.34.

Tabel 5.34 Hasil pengujian *dataset wine* dengan ID3

	true Y1	true Y2	true Y3
pred. Y1	59	71	48
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 59$, $TN = 237$, $FP = 119$, $FN = 119$ sehingga dapat digunakan untuk mencari nilai *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* dengan menggunakan metode *decision tree* ID3 diperoleh nilai *accuracy* sebesar 33.15%, nilai *recall* sebesar 33.15%, nilai *precision* sebesar 66.57%, nilai *f-measure* sebesar 44.26% dan *total cost* sebesar 28322.

5.5.29 Pengujian Algoritma C4.5 pada *Dataset Wine*

Dataset wine yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.35.

Tabel 5.35 Hasil pengujian *dataset wine* dengan C4.5

	true Y1	true Y2	true Y3
pred. Y1	56	3	0
pred. Y2	3	65	2
pred. Y3	0	3	46

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 167$, $TN = 345$, $FP = 11$, $FN = 11$ sehingga dapat digunakan untuk mencari nilai *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* dengan menggunakan metode *decision tree* C4.5 diperoleh nilai *accuracy* sebesar 93.82%, nilai *recall* sebesar

93.82%, nilai *precision* sebesar 96.91%, nilai *f-measure* sebesar 95.34% dan total cost sebesar 242.

5.5.30 Pengujian Algoritma C5.0 pada *Dataset Wine*

Dataset wine yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.36.

Tabel 5.36 Hasil pengujian *dataset wine* dengan C5.0

	true Y1	true Y2	true Y3
pred. Y1	45	10	0
pred. Y2	12	68	0
pred. Y3	0	0	2

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 203, TN = 418, FP = 12, FN = 12 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 94.42%, nilai *recall* sebesar 94.42%, nilai *precision* sebesar 97.21%, nilai *f-measure* sebesar 95.79% dan total *cost* sebesar 288.

5.5.31 Pengujian Algoritma ID3 dan PSO pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.37.

Tabel 5.37 Hasil pengujian *dataset wine* dengan ID3 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	59	71	48
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 59$, $TN = 237$, $FP = 119$, $FN = 119$ sehingga dapat digunakan untuk mencari nilai *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 33.15% ,nilai *recall* sebesar 33.15%, nilai *precision* sebesar 66.57%, nilai *f-measure* sebesar 44.26% dan total *cost* sebesar 28322.

5.5.32 Pengujian Algoritma C4.5 dan PSO pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.38.

Tabel 5.38 Hasil pengujian *dataset wine* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	56	1	0
pred. Y2	1	62	4
pred. Y3	0	4	40

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 158$, $TN = 326$, $FP = 10$, $FN = 10$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 94.05%, nilai *recall* sebesar 94.05%, nilai *precision* sebesar 97.02%, nilai *f-measure* sebesar 95.51% dan total *cost* sebesar 200.

5.5.33 Pengujian Algoritma C5.0 dan PSO pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.39.

Tabel 5.39 Hasil pengujian *dataset wine* dengan C5.0 dan PSO

	true Y1	true Y2	true Y3
pred. Y1	55	1	0
pred. Y2	2	61	4
pred. Y3	0	5	40

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 156, TN = 324, FP = 12, FN = 12 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode decision tree C5.0 diperoleh nilai *accuracy* sebesar 95.83%, nilai *recall* sebesar 95.83%, nilai *precision* sebesar 97.92%, nilai *f-measure* sebesar 94.61% dan total *cost* sebesar 288.

5.5.34 Pengujian *Cost Sensitive Decision Tree ID3* pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.40.

Tabel 5.40 Hasil pengujian *dataset wine* dengan csDT ID3

	true Y1	true Y2	true Y3
pred. Y1	59	71	48
pred. Y2	0	0	0
pred. Y3	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 59, TN = 237, FP = 119, FN = 119 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) ID3 diperoleh nilai *accuracy* sebesar 33.15%, nilai *recall* sebesar 33.15%, nilai *precision* sebesar 66.57%, nilai *f-measure* sebesar 44.26% dan total *cost* sebesar 28322.

5.5.35 Pengujian *Cost Sensitive Decision Tree C4.5* pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan

algoritma C4.5 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.41.

Tabel 5.41 Hasil pengujian *dataset wine* dengan csDT C4.5

	true Y1	true Y2	true Y3
pred. Y1	56	1	0
pred. Y2	1	64	0
pred. Y3	0	2	44

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 164, TN = 332, FP = 4, FN = 4 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) C4.5 diperoleh nilai *accuracy* sebesar 97.62%, nilai *recall* sebesar 97.62%, nilai *precision* sebesar 98.81%, nilai *f-measure* sebesar 98.21% dan total *cost* sebesar 32.

5.5.36 Pengujian *Cost Sensitive Decision Tree C5.0* pada *Dataset Wine*

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.42.

Tabel 5.42 Hasil pengujian *dataset wine* dengan csDT C5.0

	true Y1	true Y2	true Y3
pred. Y1	56	1	0
pred. Y2	1	64	3
pred. Y3	0	2	41

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 161, TN = 329, FP = 7, FN = 7 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) C5.0 diperoleh nilai *accuracy* sebesar 95.83%, nilai *recall* sebesar 95.83%, nilai *precision* sebesar 97.92%, nilai *f-measure* sebesar 96.86% dan total *cost* sebesar 98.

5.5.37 Pengujian Algoritma ID3 pada Data *Vehicle*

Dataset vehicle yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.43.

Tabel 5.43 Hasil pengujian *dataset vehicle* dengan ID3

	true Y1	true Y2	true Y3	true Y4
pred. Y1	96	88	108	84
pred. Y2	0	0	0	0
pred. Y3	0	0	0	0
pred. Y4	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 96, TN = 848, FP = 280, FN = 280 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 25.53%, nilai *recall* sebesar 25.53%, nilai *precision* sebesar 75.18%, nilai *f-measure* sebesar 38.12% dan total *cost* sebesar 156800.

5.5.38 Pengujian Algoritma C4.5 pada Data *Vehicle*

Dataset vehicle yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.44.

Tabel 5.44 Hasil pengujian *dataset vehicle* dengan C4.5

	true Y1	true Y2	true Y3	true Y4
pred. Y1	79	7	3	5
pred. Y2	11	47	2	34
pred. Y3	2	11	101	5
pred. Y4	4	23	2	40

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 267, TN = 1019, FP = 109, FN = 109 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle*

dengan menggunakan metode decision tree C4.5 diperoleh nilai accuracy sebesar 71.01%, nilai *recall* sebesar 71.01%, nilai *precision* sebesar 90.34%, nilai *f-measure* sebesar 79.54% dan total *cost* sebesar 23762.

5.5.39 Pengujian Algoritma C5.0 pada Data Vehicle

Dataset vehicle yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 1. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.45.

Tabel 5.45 Hasil pengujian *dataset vehicle* dengan C5.0

	true Y1	true Y2	true Y3	true Y4
pred. Y1	85	6	4	5
pred. Y2	4	45	2	30
pred. Y3	2	6	97	7
pred. Y4	5	31	5	42

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 269, TN = 1021, FP = 107, FN = 107 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* dengan menggunakan metode decision tree C5.0 diperoleh nilai accuracy sebesar 71.54%, nilai *recall* sebesar 71.54%, nilai *precision* sebesar 90.51%, nilai *f-measure* sebesar 90.51% dan total *cost* sebesar 22898.

5.5.40 Pengujian Algoritma ID3 dan PSO pada Data Vehicle

Dataset vehicle yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma ID3 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.46.

Tabel 5.46 Hasil pengujian *dataset vehicle* dengan ID3 dan PSO

	true Y1	true Y2	true Y3	true Y4
pred. Y1	96	88	108	84
pred. Y2	0	0	0	0
pred. Y3	0	0	0	0
pred. Y4	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 96$, $TN = 848$, $FP = 280$, $FN = 280$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode decision tree ID3 diperoleh nilai *accuracy* sebesar 25.53%, nilai *recall* sebesar 25.53%, nilai *precision* sebesar 75.18%, nilai *f-measure* sebesar 38.12% dan total *cost* sebesar 156800.

5.5.41 Pengujian Algoritma C4.5 dan PSO pada Data Vehicle

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.47.

Tabel 5.47 Hasil pengujian *dataset vehicle* dengan C4.5 dan PSO

	true Y1	true Y2	true Y3	true Y4
pred. Y1	84	4	2	5
pred. Y2	3	49	5	27
pred. Y3	2	7	93	6
pred. Y4	7	28	8	46

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 272$, $TN = 1024$, $FP = 104$, $FN = 104$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode decision tree C4.5 diperoleh nilai *accuracy* sebesar 72.34%, nilai *recall* sebesar 72.34%, nilai *precision* sebesar 90.78%, nilai *f-measure* sebesar 80.52% dan total *cost* sebesar 21632.

5.5.42 Pengujian Algoritma C5.0 dan PSO pada Data Vehicle

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 2. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.48.

Tabel 5.48 Hasil pengujian *dataset vehicle* dengan C5.0 dan PSO

	true Y1	true Y2	true Y3	true Y4
pred. Y1	82	7	2	3
pred. Y2	6	46	4	29
pred. Y3	1	6	95	8
pred. Y4	7	29	7	44

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 267$, $TN = 1019$, $FP = 109$, $FN = 109$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 71.01%, nilai *recall* sebesar 71.01%, nilai *precision* sebesar 90.34%, nilai *f-measure* sebesar 79.52% dan *total cost* sebesar 23762.

5.5.43 Pengujian *Cost Sensitive Decision Tree* ID3 pada Data *Vehicle*

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive decision tree* (csDT) ID3 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.49.

Tabel 5.49 Hasil pengujian *dataset vehicle* dengan csDT ID3

	true Y1	true Y2	true Y3	true Y4
pred. Y1	96	88	108	84
pred. Y2	0	0	0	0
pred. Y3	0	0	0	0
pred. Y4	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 96$, $TN = 848$, $FP = 280$, $FN = 280$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) ID3 diperoleh nilai *accuracy* sebesar 25.53%, nilai *recall* sebesar 25.53%, nilai *precision* sebesar 75.18%, nilai *f-measure* sebesar 38.12% dan *total cost* sebesar 156800.

5.5.44 Pengujian *Cost Sensitive Decision Tree C4.5* pada Data *Vehicle*

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C4.5 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.50.

Tabel 5.50 Hasil pengujian *dataset vehicle* dengan csDT C4.5

	true Y1	true Y2	true Y3	true Y4
pred. Y1	90	6	2	3
pred. Y2	4	54	5	29
pred. Y3	1	7	100	7
pred. Y4	1	21	1	45

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 289, TN = 1041, FP = 87, FN = 87 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) C4.5 diperoleh nilai *accuracy* sebesar 76.86%, nilai *recall* sebesar 76.86%, nilai *precision* sebesar 92.29%, nilai *f-measure* sebesar 83.87% dan total *cost* sebesar 15138.

5.5.45 Pengujian *Cost Sensitive Decision Tree C5.0* pada Data *Vehicle*

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 376 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan algoritma C5.0 pada perancangan skenario 3. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.51.

Tabel 5.51 Hasil pengujian *dataset vehicle* dengan csDT C5.0

	true Y1	true Y2	true Y3	true Y4
pred. Y1	88	6	2	3
pred. Y2	5	50	2	27
pred. Y3	1	6	101	10
pred. Y4	2	26	3	44

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 283, TN = 1035, FP = 93, FN = 93 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle*

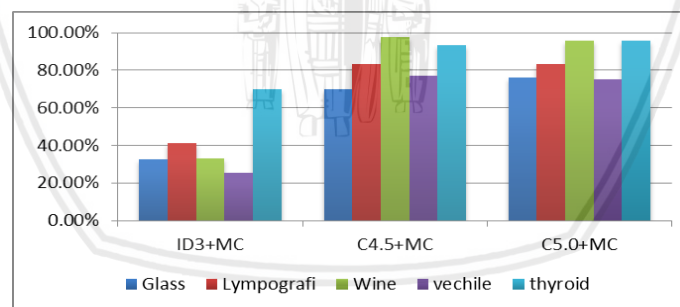
PSO dengan menggunakan metode *cost sensitive decision tree* (csDT) C5.0 diperoleh nilai *accuracy* sebesar 75.27%, nilai *recall* sebesar 75.27%, nilai *precision* sebesar 91.76%, nilai *f-measure* sebesar 82.70% dan total *cost* sebesar 17298.

5.6 Pembahasan Pengujian Metode *Decision Tree*

Hasil dari pengujian metode *decision tree* dengan mengimplementasikan algoritma ID3, C4.5 dan C5.0 dengan menggunakan skenario 1, skenario 2 dan skenario 3 pada *dataset glass*, *lympografi*, *vehicle*, *thyroid* dan *wine* diperoleh hasil dari pengujian berupa nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Nilai parameter dari pengujian metode tersebut dapat dianalisa sebagai berikut.

Tabel 5.52 Nilai Parameter *Accuracy* ID3, C4.5 dan C5.0

DATASET	ID3			C4.5			C5.0		
	ID3	ID3+ PSO	ID3+PSO +META	C4.5	C4.5+ PSO	C4.5+PSO +META	C5.0	C5.0+ PSO	C5.0+PSO +META
Vehicle	25.53%	25.53%	25.53%	71.01%	72.34%	76.86%	71.54%	71.54%	75.27%
Glass	32.71%	32.71%	32.71%	68.22%	68.22%	70.09%	67.29%	70.09%	76.17%
Lympografi	41.22%	41.22%	41.22%	75.68%	79.71%	83.33%	75.00%	78.99%	83.33%
Wine	33.15%	33.15%	33.15%	93.82%	94.05%	97.62%	94.38%	94.38%	95.83%
Thyroid	69.77%	69.77%	69.77%	93.95%	93.95%	93.49%	94.42%	94.42%	95.81%



Gambar 5.11 Diagram Nilai *Accuracy Dataset*

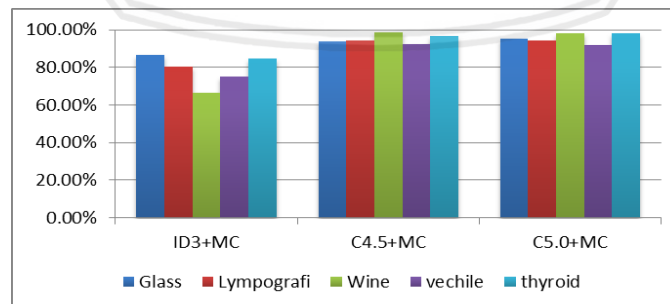
Parameter *accuracy* menunjukkan tingkat kedekatan antara nilai prediksi dengan nilai aktual jadi semakin besar prosentase nilai *accuracy* maka tingkat ketepatan prediksi semakin baik. Berdasarkan Tabel 5.52 hasil pengujian diperoleh nilai *accuracy* yang digunakan untuk mengukur performa metode dalam mengklasifikasikan *dataset*. Pada Tabel 5.52 menunjukkan bahwa algoritma ID3 memiliki nilai *accuracy* yang kecil karena tidak mampu mengklasifikasikan semua *dataset* uji yang memiliki tipe data *real* dan *interger*. Pada

algoritma C4.5 dan C5.0 dengan menggunakan *particle swarm optimization (PSO)* untuk menyeleksi atribut. Pada Gambar 5.6 hasil yang diperoleh pada algoritma C4.5 dengan menggunakan PSO pada *dataset vehicle, lypnografi, wine* dapat meningkat masing-masing 72.34%, 79.71%, 94.05%.

Sedangkan hasil dari algoritma C5.0 dengan menggunakan PSO pada *dataset glass, lypnografi* mampu meningkatkan nilai *accuracy* masing-masing 70.09%, 78.99%. Hal tersebut membuktikan bahwa pemilihan atribut memiliki pengaruh pada performa *decision tree*. Hasil pengujian dengan menggunakan menggunakan metode *cost sensitive decision tree* pada ID3, C4.5 dan C5.0 dimana dengan menggunakan metode *cost sensitive decision tree* C4.5 pada *dataset vehicle* dan *wine* memiliki nilai *accuracy* yang baik masing-masing sebesar 76.86% dan 97.62% dari pada menggunakan C5.0. Pada *dataset glass* dan *thyroid* memiliki nilai *accuracy* yang baik masing-masing 76.17% dan 95.81% jika menggunakan *cost sensitive decision tree* C5.0.

Tabel 5.53 Nilai Parameter *Recall* ID3, C4.5 dan C5.0

DATASET	ID3			C4.5			C5.0		
	ID3	ID3+ PSO	ID3+PSO +META	C4.5	C4.5+ PSO	C4.5+PSO +META	C5.0	C5.0+ PSO	C5.0+PSO +META
Vehicle	25.53%	25.53%	25.53%	71.01%	72.34%	76.86%	71.54%	71.01%	75.27%
Glass	32.71%	32.71%	32.71%	68.22%	68.22%	70.09%	67.29%	70.09%	76.17%
Lypnografi	41.22%	41.22%	41.22%	75.68%	79.71%	83.33%	75.00%	78.99%	83.33%
Wine	33.15%	33.15%	33.15%	93.82%	94.05%	97.62%	94.38%	92.86%	95.83%
Thyroid	69.77%	69.77%	69.77%	93.95%	93.95%	93.44%	94.42%	94.42%	95.81%

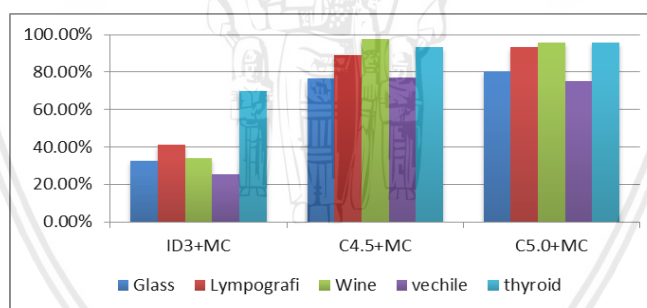


Gambar 5.12 Diagram Nilai *Recall* Dataset

Parameter *recall* menunjukkan tingkat keberhasilan sistem dalam menemukan informasi jadi semakin besar nilai prosentase maka data pengujian mampu mengenali pola *decision tree* yang telah dilatih dengan baik. Berdasarkan Tabel 5.53 menunjukkan bahwa hasil nilai *recall* pada *dataset vehicle* dan *wine* diperoleh hasil terbaik masing –masing sebesar 76.86% dan 97.62% dengan menggunakan metode *cost sensitive decision tree* C4.5 sedangkan pada *dataset glass* dan *thyroid* diperoleh nilai terbaik dengan menggunakan metode *cost sensitive decision tree* C5.0 dengan masing-masing nilai *recall* sebesar 76.17% dan 95.81%.

Tabel 5.54 Nilai Parameter *precision* ID3, C4.5 dan C5.0

DATASET	ID3			C4.5			C5.0		
	ID3	ID3+ PSO	ID3+PSO +META	C4.5	C4.5+ PSO	C4.5+PSO +META	C5.0	C5.0+ PSO	C5.0+PSO +META
Vehicle	75.18%	75.18%	75.18%	90.34%	90.78%	92.29%	90.51%	90.34%	91.76%
Glass	86.54%	86.54%	86.54%	93.64%	93.64%	94.02%	93.46%	94.02%	95.23%
Lymphografi	80.41%	80.41%	80.41%	91.89%	93.24%	94.44%	91.67%	93.00%	94.44%
Wine	66.57%	66.57%	66.57%	96.91%	97.02%	98.81%	97.19%	96.43%	97.92%
Thyroid	84.88%	84.88%	84.88%	96.98%	96.98%	96.74%	97.21%	97.21%	97.91%

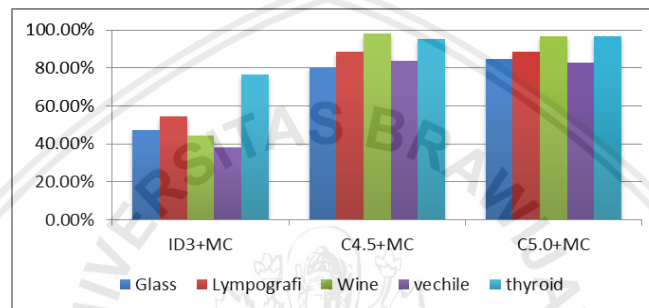


Gambar 5.13 Diagram Nilai *Precision* Dataset

Parameter *precision* menunjukkan tingkat ketepatan antara informasi data yang diminta dengan hasil prediksi jadi semakin besar nilai prosentase maka informasi yang diinginkan sesuai dengan hasil prediksi. Berdasarkan Tabel 5.54 menunjukkan bahwa hasil nilai *precision* pada *dataset vehicle* dan *wine* diperoleh hasil terbaik masing – masing sebesar 92.29% dan 98.81% dengan menggunakan metode *cost sensitive decision tree* C4.5 sedangkan pada *dataset glass* dan *thyroid* diperoleh nilai terbaik dengan menggunakan metode *cost sensitive decision tree* C5.0 dengan masing-masing nilai *precision* sebesar 95.23% dan 97.91%.

Tabel 5.55 Nilai Parameter F-Measure ID3, C4.5 dan C5.0

DATASET	ID3			C4.5			C5.0		
	ID3	ID3+ PSO	ID3+PSO +META	C4.5	C4.5+ PSO	C4.5+PSO +META	C5.0	C5.0+ PSO	C5.0+PSO +META
Vehicle	38.12%	38.12%	38.12%	79.52%	80.52%	83.87%	79.92%	90.34%	82.70%
Glass	47.48%	47.48%	47.48%	78.94%	78.94%	80.31%	78.24%	94.02%	84.64%
Lympografi	54.50%	54.50%	54.50%	83.00%	85.94%	88.54%	82.50%	93.00%	88.54%
Wine	44.26%	44.26%	44.26%	95.34%	95.51%	98.21%	95.77%	96.43%	96.86%
Thyroid	76.59%	76.59%	76.59%	95.44%	95.44%	95.09%	95.79%	97.21%	96.85%

**Gambar 5.14** Diagram Nilai *F-Measure* Dataset

Parameter *F-Measure* menunjukkan tingkat performa prediksi pada *decision tree* dengan menemukan informasi dan ketepatan prediksi informasi pada *dataset*. Pada Tabel 5.55 menunjukkan bahwa hasil nilai *f-measure* pada *dataset vehicle dan wine* diperoleh hasil terbaik masing –masing sebesar 83.87% dan 98.21% dengan menggunakan metode *cost sensitive decision tree* C4.5 sedangkan pada *dataset glass dan thyroid* diperoleh nilai terbaik dengan menggunakan metode *cost sensitive decision tree* C5.0 dengan masing-masing nilai *recall* sebesar 84.64% dan 96.85%.

Tabel 5.56 Nilai Total Cost ID3, C4.5 dan C5.0

DATASET	ID3			C4.5			C5.0		
	ID3	ID3+ PSO	ID3+PSO +META	C4.5	C4.5+ PSO	C4.5+PSO +META	C5.0	C5.0+ PSO	C5.0+PSO +META
Vehicle	156800	156800	156800	23762	21632	15138	22898	23762	17298
Glass	41472	41472	41472	9248	9248	8192	9800	8192	5202
Lympografi	15138	15138	15138	2592	1568	1058	2738	1682	1058
Wine	28322	28322	28322	242	200	32	288	200	98
Thyroid	8450	8450	8450	338	338	392	288	288	162

Parameter total *cost* menunjukkan bahwa semakin besar nilai *cost* yang dihasilkan maka tingkat kesalahan prediksi data semakin besar. Pada Tabel 5.56 menunjukkan bahwa hasil nilai total *cost* pada *dataset vehicle dan wine* diperoleh *cost* terkecil masing – masing sebesar 15138 dan 32 dengan menggunakan metode *cost sensitive decision tree* C4.5 sedangkan pada *dataset glass dan thyroid* diperoleh nilai terbaik dengan menggunakan metode *cost sensitive decision tree* C5.0 dengan masing-masing nilai *recall* sebesar 5202 dan 162.

5.7 Pengujian Cost Sensitive Naïve Bayes

Pada penelitian ini dilakukan pengujian menggunakan metode klasifikasi yang memiliki karakteristik yang hampir sama seperti metode *decision tree*. Dalam pengujian ini metode *cost sensitive learning* digunakan untuk mengoptimalkan metode *naïve bayes*. Tujuan penelitian ini untuk membandingkan hasil dari performa *cost sensitive decision tree* dengan *cost sensitive naïve bayes* dalam mengklasifikasikan data *imbalanced multiclass*. Pengujian dilakukan pada lima *dataset* dari *UCI machine learning repository*. *Dataset* yang digunakan dalam pengujian terdiri atas *dataset multiclass* antara lain *vehicle, glass, lymfografi, thyroid dan wine*.

5.7.1 Pengujian Cost Sensitive Naïve Bayes pada Dataset Glass

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.57.

Tabel 5.57 Hasil pengujian *dataset glass* dengan csNB

	Actual Y1	Actual Y2	Actual Y3	Actual Y5	Actual Y6	Actual Y7
Pred. Y1	70	76	17	13	9	29
Pred. Y2	0	0	0	0	0	0
Pred. Y3	0	0	0	0	0	0
Pred. Y5	0	0	0	0	0	0
Pred. Y6	0	0	0	0	0	0
Pred. Y7	0	0	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 70$, $TN = 926$, $FP = 144$, $FN = 144$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 32.71%, nilai *recall* sebesar 32.71%, nilai *precision* sebesar 86.54%, nilai *f-measure* sebesar 47.48% dan *total cost* sebesar 41472.

5.7.2 Pengujian Cost Sensitive Naïve Bayes pada Dataset Glass PSO

Dataset glass PSO yang digunakan dalam pengujian ini memiliki data sebanyak 214 dengan jumlah *class* sebanyak 6 dan 10 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.58.

Tabel 5.58 Hasil pengujian *dataset glass* PSO dengan csNB

	true Y1	true Y2	true Y3	true Y5	true Y6	true Y7
pred. Y1	70	76	17	13	9	29
pred. Y2	0	0	0	0	0	0
pred. Y3	0	0	0	0	0	0
pred. Y5	0	0	0	0	0	0
pred. Y6	0	0	0	0	0	0
pred. Y7	0	0	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 70$, $TN = 926$, $FP = 144$, $FN = 144$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset glass* PSO dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 32.71%, nilai *recall* sebesar 32.71%, nilai *precision* sebesar 86.54%, nilai *f-measure* sebesar 47.48% dan *total cost* sebesar 41472.

5.7.3 Pengujian Cost Sensitive Naïve Bayes pada Dataset Lymphografi

Dataset lymphografi yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.59.

Tabel 5.59 Hasil pengujian *dataset lypografi* dengan csNB

	true Y1	true Y2	true Y3	true Y5
pred. Y1	41	14	1	1
pred. Y2	19	67	2	1
pred. Y3	1	0	1	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 109$, $TN = 405$, $FP = 39$, $FN = 39$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 73.65%, nilai *recall* sebesar 73.65%, nilai *precision* sebesar 91.22%, nilai *f-measure* sebesar 81.50% dan *total cost* sebesar 3042.

5.7.4 Pengujian *Cost Sensitive Naïve Bayes* pada *Dataset Lypografi* PSO

Dataset glass yang digunakan dalam pengujian ini memiliki data sebanyak 148 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.60.

Tabel 5.60 Hasil pengujian *dataset lypografi* PSO dengan csNB

	true Y1	true Y2	true Y3	true Y5
pred. Y1	44	8	2	0
pred. Y2	13	70	0	1
pred. Y3	0	0	0	0
pred. Y5	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 114$, $TN = 390$, $FP = 24$, $FN = 24$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset lypografi* dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy*

sebesar 82.61%, nilai *recall* sebesar 82.61%, nilai *precision* sebesar 94.20%, nilai *f-measure* sebesar 88.03% dan total *cost* sebesar 1152.

5.7.5 Pengujian Cost Sensitive Naïve Bayes pada Dataset Thyroid

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.61.

Tabel 5.61 Hasil pengujian *dataset thyroid* dengan csNB

	true Y1	true Y2	true Y3
pred. Y1	146	0	4
pred. Y2	0	35	0
pred. Y3	4	0	26

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 207, TN = 422, FP = 8, FN = 8 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 96.28%, nilai *recall* sebesar 96.28%, nilai *precision* sebesar 98.14%, nilai *f-measure* sebesar 97.20% dan total *cost* sebesar 128.

5.7.6 Pengujian Cost Sensitive Naïve Bayes pada Dataset Thyroid PSO

Dataset thyroid PSO yang digunakan dalam pengujian ini memiliki data sebanyak 215 dengan jumlah *class* sebanyak 3 dan 4 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.62.

Tabel 5.62 Hasil pengujian *dataset thyroid* PSO dengan csNB

	true Y1	true Y2	true Y3
pred. Y1	147	1	1
pred. Y2	0	34	0
pred. Y3	3	0	29

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 210$, $TN = 425$, $FP = 5$, $FN = 5$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset thyroid* PSO dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 97.67%, nilai *recall* sebesar 97.67%, nilai *precision* sebesar 98.84%, nilai *f-measure* sebesar 98.25% dan *total cost* sebesar 50.

5.7.7 Pengujian Cost Sensitive Naïve Bayes pada Dataset Wine

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.63.

Tabel 5.63 Hasil pengujian *dataset wine* dengan csNB

	true Y1	true Y2	true Y3
pred. Y1	57	4	0
pred. Y2	2	65	1
pred. Y3	0	2	47

Berdasarkan tabel *confusion matrix* diperoleh nilai $TP = 169$, $TN = 347$, $FP = 9$, $FN = 9$ sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan *total cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 94.94%, nilai *recall* sebesar 94.94%, nilai *precision* sebesar 97.47%, nilai *f-measure* sebesar 96.19% dan *total cost* sebesar 162.

5.7.8 Pengujian Cost Sensitive Naïve Bayes pada Dataset Wine PSO

Dataset wine PSO yang digunakan dalam pengujian ini memiliki data sebanyak 178 dengan jumlah *class* sebanyak 3 dan 13 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.64.

Tabel 5.64 Hasil pengujian *dataset wine* PSO dengan csNB

	true Y1	true Y2	true Y3
pred. Y1	57	4	0
pred. Y2	2	65	1
pred. Y3	0	2	47

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 169, TN = 347, FP = 9, FN = 9 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset wine* PSO dengan menggunakan metode *cost sensitive naïve bayes* diperoleh nilai *accuracy* sebesar 94.94%, nilai *recall* sebesar 94.94%, nilai *precision* sebesar 97.47%, nilai *f-measure* sebesar 96.19% dan total *cost* sebesar 162.

5.7.9 Pengujian *Cost Sensitive Naïve Bayes* pada *Dataset Vehicle*

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 946 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti Tabel 5.65.

Tabel 5.65 Hasil pengujian *dataset vehicle* dengan csNB

	true Y1	true Y2	true Y3	true Y4
pred. Y1	96	88	108	84
pred. Y2	0	0	0	0
pred. Y3	0	0	0	0
pred. Y4	0	0	0	0

Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 96, TN = 848, FP = 280, FN = 280 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 25.53%, nilai *recall* sebesar 25.53%, nilai *precision* sebesar 75.18%, nilai *f-measure* sebesar 38.12% dan total *cost* sebesar 156800.

5.7.10 Pengujian *Cost Sensitive Naïve Bayes* pada *Dataset Vehicle PSO*

Dataset vehicle PSO yang digunakan dalam pengujian ini memiliki data sebanyak 946 dengan jumlah *class* sebanyak 4 dan 18 atribut. Pengujian ini dilakukan dengan menggunakan *cost sensitive naïve bayes* (csNB) pada perancangan skenario 4. Hasil dari pengujian tersebut berupa tabel *confusion matrix* seperti tabel 5.66.

Tabel 5.66 Hasil pengujian *dataset vehicle* PSO dengan csNB

	true Y1	true Y2	true Y3	true Y4
pred. Y1	96	88	108	84
pred. Y2	0	0	0	0
pred. Y3	0	0	0	0
pred. Y4	0	0	0	0

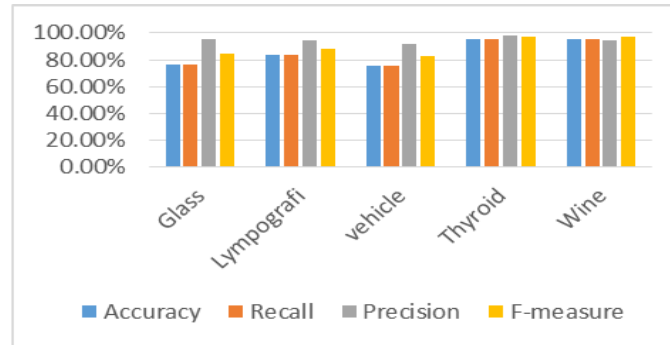
Berdasarkan tabel *confusion matrix* diperoleh nilai TP = 96, TN = 848, FP = 280, FN = 280 sehingga dapat digunakan untuk mencari nilai parameter *accuracy*, *precision*, *recall*, *f-measure* dan total *cost*. Berdasarkan hasil perhitungan nilai parameter pada *dataset vehicle* PSO dengan menggunakan metode *decision tree* C5.0 diperoleh nilai *accuracy* sebesar 25.53%, nilai *recall* sebesar 25.53%, nilai *precision* sebesar 75.18%, nilai *f-measure* sebesar 38.12% dan total *cost* sebesar 156800.

5.8 Pembahasan Perbandingan Pengujian csDT dan csNB

Pengujian dalam penelitian dilakukan dengan dua metode, yaitu dengan menggunakan metode *Cost Sensitive Decision Tree* C5.0 (csDT) dan metode *Cost Sensitive Naive Bayes* (csNB). Pada pengujian pertama ini percobaan dilakukan dengan menguji lima *dataset* yang telah dilakukan seleksi atribut dengan PSO menggunakan metode csDT C5.0. Hasil dari pengujian ditampilkan sebagai berikut:

Tabel 5.67 Hasil Pengujian csDT C5.0

<i>Dataset</i>	<i>Accuracy</i>	<i>Recall</i>	<i>Precision</i>	<i>F-measure</i>
<i>Glass</i>	76.17%	76.17%	95.23%	84.64%
<i>Lymphografi</i>	83.33%	83.33%	94.44%	88.54%
<i>Vehicle</i>	75.27%	75.27%	91.76%	82.70%
<i>Thyroid</i>	95.81%	95.81%	97.91%	96.85%
<i>Wine</i>	95.83%	95.83%	94.44%	96.86%



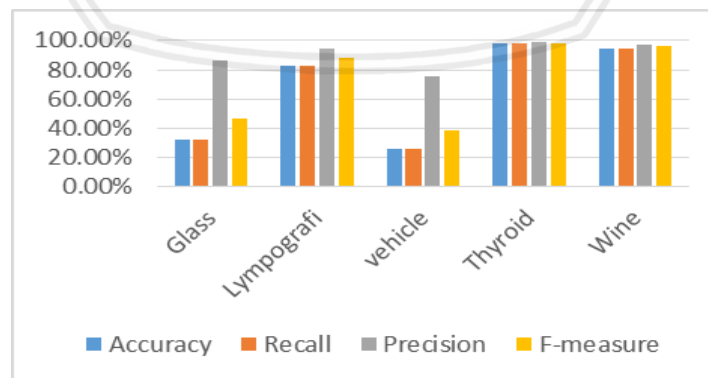
Gambar 5.15 Diagram Hasil Pengujian csDT

Berdasarkan Tabel 5.67 menampilkan nilai parameter *accuracy* pada *dataset glass*, *lympografi*, *vehicle*, *thyroid* dan *wine* masing-masing 76.17%, 83.33%, 75.27%, 95.81% dan 95.83%. Gambar 5.15 menggambarkan hasil pengukuran keempat parameter yaitu *accuracy*, *recall*, *precision* dan *f-measure*. Gambar tersebut menjelaskan bahwa kelima dataset memiliki hasil klasifikasi yang cukup baik dengan menggunakan metode csDT.

Pada pengujian kedua percobaan dilakukan dengan menggunakan metode csNB pada kelima *dataset* yang telah dilakukan seleksi atribut dengan PSO. Hasil dari pengujian ditampilkan sebagai berikut:

Tabel 5.68 Hasil Pengujian csNB

<i>Dataset</i>	<i>Accuracy</i>	<i>Recall</i>	<i>Precision</i>	<i>F-measure</i>
<i>Glass</i>	32.24%	32.24%	86.45%	46.97%
<i>Lymphografi</i>	82.61%	82.61%	94.20%	88.03%
<i>vehicle</i>	25.53%	25.53%	75.18%	38.12%
<i>Thyroid</i>	97.67%	97.67%	98.84%	98.25%
<i>Wine</i>	94.94%	94.94%	97.47%	96.19%



Gambar 5.16 Diagram Hasil Pengujian csNB

Tabel 5.68 merupakan hasil pengujian menggunakan metode csNB pada lima *dataset* yaitu *glass*, *lympografi*, *vehicle*, *thyroid*, *wine* masing-masing *dataset* memiliki nilai *accuracy* sebagai berikut 32.24%, 82.61%, 25.53%, 97.67%, 94.94%. Gambar 5.16 menunjukkan *dataset thyroid* dan *wine* memiliki nilai parameter yang baik. Namun, pada *dataset glass* dan *vehicle* memiliki hasil parameter yang kurang. Hal tersebut mengindikasikan bahwa metode csNB kurang baik untuk mengklasifikasikan pada tipe data *glass* dan *vehicle*.

Hasil dari pengujian menggunakan metode *cost sensitive naïve bayes* dengan menggunakan skenario 4 pada *dataset glass*, *lympografi*, *vehicle*, *thyroid* dan *wine* diperoleh hasil dari pengujian berupa nilai parameter *accuracy*, *recall*, *precision*, *f-measure* dan total *cost*. Kemudian hasil pengujian tersebut dibandingkan dengan metode *cost sensitive decision tree C5.0*. Diperoleh hasil perbandingan kedua metode tersebut sebagai berikut.

Tabel 5.69 Perbandingan Nilai Accuracy csDT dan csNB

DATASET	csDT		csNB	
	csDT	CsDT +PSO	csNB	CsNB +PSO
Vehicle	67.29%	75.27%	25.53%	25.53%
Glass	75.00%	76.17%	32.71%	32.71%
Lympografi	71.01%	83.33%	73.65%	82.61%
Wine	94.42%	95.83%	94.94%	94.94%
Thyroid	92.86%	95.81%	96.28%	97.67%

Berdasarkan Tabel 5.69 membuktikan bahwa dengan menggunakan data hasil seleksi PSO mampu menghasilkan performa *accuracy* yang lebih baik pada kedua *classifier*. Hasil pengujian pada *dataset vehicle*, *glass*, *lympografi* dan *wine* memiliki performa *accuracy* yang baik dengan menggunakan metode *cost sensitive decision tree C5.0* dibandingkan dengan menggunakan *cost sensitive naïve bayes* pada *dataset vehicle*, *glass*, *lympografi* dan *wine* masing-masing nilai *accuracy* pengujian sebesar 75.27%, 76.17%, 83.33%, 95.83%. sedangkan pada *dataset thyroid* memiliki nilai *accuracy* yang baik dengan menggunakan metode *cost sensitive naïve bayes* dengan nilai *accuracy* sebesar 97.67%.

Tabel 5.70 Hasil Perhitungan *Cost*

<i>Dataset</i>	<i>Decision Tree</i>			<i>Naïve Bayes</i>		
	<i>DT</i>	<i>DT + PSO</i>	<i>csDT</i>	<i>NB</i>	<i>NB + PSO</i>	<i>csNB</i>
<i>Glass</i>	9800	8192	5202	40898	40898	42050
<i>Lymphografi</i>	2738	1682	1058	2888	1458	1152
<i>vehicle</i>	22898	23762	17298	95048	95048	156800
<i>Thyroid</i>	288	288	162	98	50	50
<i>Wine</i>	288	200	98	18	18	162

Tabel 5.70 merupakan hasil perhitungan *cost* dari *classifier*. *Cost* adalah nilai kesalahan dari klasifikasi, jadi semakin besar nilai *cost* maka *classifier* memiliki performa yang buruk. Hasil pengujian dengan menggunakan csDT mampu meminimalkan *cost* pada semua *dataset*. Sedangkan pengujian menggunakan csNB hanya mampu meminimalkan *cost* pada *dataset lymphografi*.

