

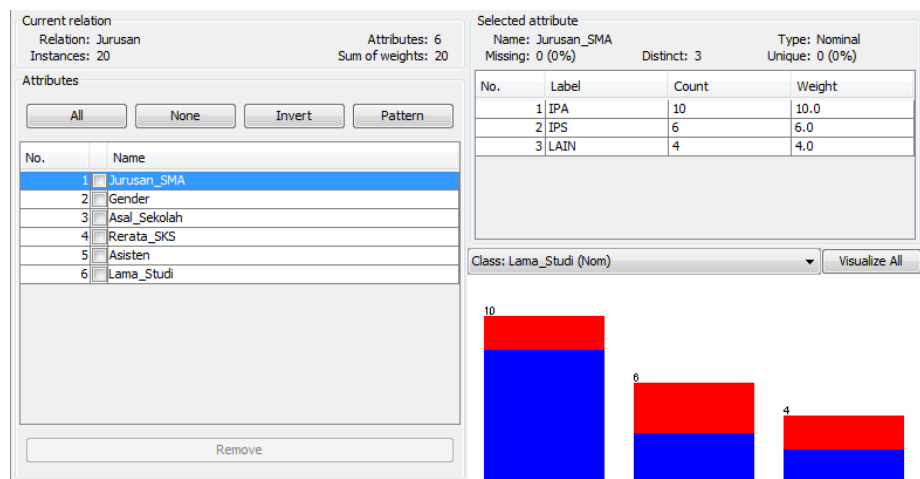
## MODUL 9

### TUGAS

1. Berdasarkan pohon keputusan pada kegiatan 9.4.2 (menggunakan RapidMiner), isikan nilai kelas atribut Bermain\_Tenis pada tabel Testing berikut:

Cuaca	Suhu	Kelembaban_Udara	Berangin	Bermain_Tenis
Cerah	75	65	TIDAK	YA
Cerah	80	68	YA	YA
Cerah	83	87	YA	TIDAK
Mendung	70	96	TIDAK	YA
Mendung	68	81	TIDAK	YA
Hujan	65	75	TIDAK	YA
Hujan	64	85	YA	TIDAK

2. Gunakan file ARFF yang dikerjakan pada tugas nomor 1 dalam Modul 7 sebagai data training
  - a. Buatlah dan cetaklah pohon keputusan berdasarkan data tersebut



```

Classifier output

=== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Jurusan
Instances:   20
Attributes:  6
             Jurusan_SMA
             Gender
             Asal_Sekolah
             Rerata_SKS
             Asisten
             Lama_Studi

Test mode:   evaluate on training data

=== Classifier model (full training set) ===

J48 pruned tree

Gender = WANITA: TEPAT (5.0/1.0)
Gender = PRIA
| Rerata_SKS <= 18: TERLAMBAT (4.0)
| Rerata_SKS > 18: TEPAT (11.0/2.0)

Number of Leaves :    3

Size of the tree :    5

Time taken to build model: 0 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

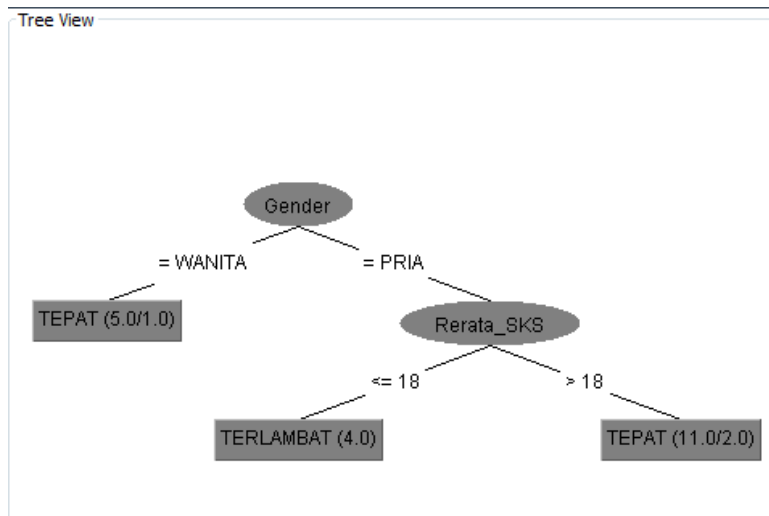
```

=== Summary ===

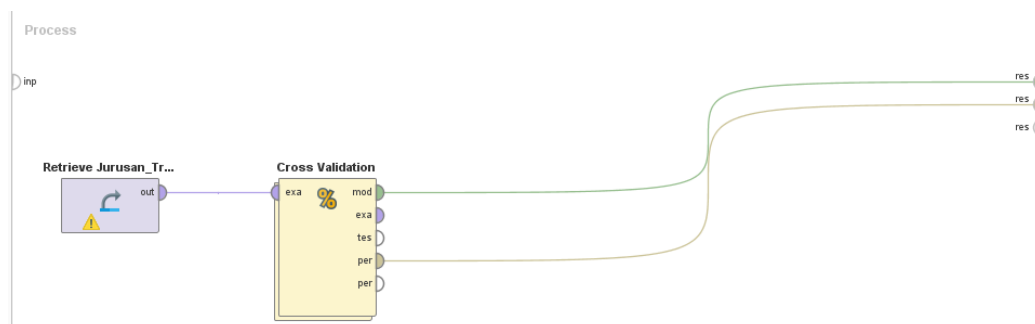
Correctly Classified Instances	17	85	%
Incorrectly Classified Instances	3	15	%
Kappa statistic	0.6341		
Mean absolute error	0.2436		
Root mean squared error	0.349		
Relative absolute error	53.0693	%	
Root relative squared error	73.1456	%	
Coverage of cases (0.95 level)	100	%	
Mean rel. region size (0.95 level)	90	%	
Total Number of Instances	20		

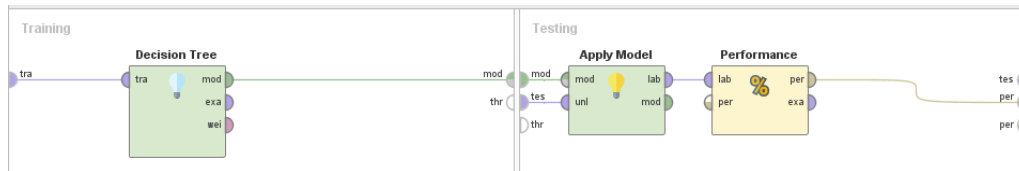
=== Confusion Matrix ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	a	b	<-- classified as
1,000	0,429	0,813	1,000	0,897	0,897	0,681	13	0	a = TEPAT
0,571	0,000	1,000	0,571	0,727	0,681	0,681	3	4	b = TERLAMBAT
Weighted Avg.	0,850	0,279	0,878	0,850	0,837	0,681			



- b. Carilah nilai-nilai berikut ini:
  - i. Jumlah simpul daun pada pohon keputusan = 3
  - ii. Jumlah simpul keseluruhan pada pohon keputusan = 5
  - iii. Waktu yang dibutuhkan untuk proses pelatihan = 0 detik
  - iv. Tingkat ketepatan klasifikasi = 85%
  - v. Tingkat ketidakpastian klasifikasi = 15%
3. Gunakan file excel yang dikerjakan pada tugas nomor 1 dalam Modul 6 sebagai data training.
  - a. Buatlah dan cetaklah pohon keputusan berdasarkan data tersebut





Parameters

Decision Tree

criterion ✓ information\_gain ⓘ

maximal depth ✓ 10 ⓘ

☒ apply pruning ⓘ

confidence ✓ 0.1 ⓘ

☒ apply prepruning ⓘ

minimal gain ✓ 0.01 ⓘ

minimal leaf size 2 ⓘ

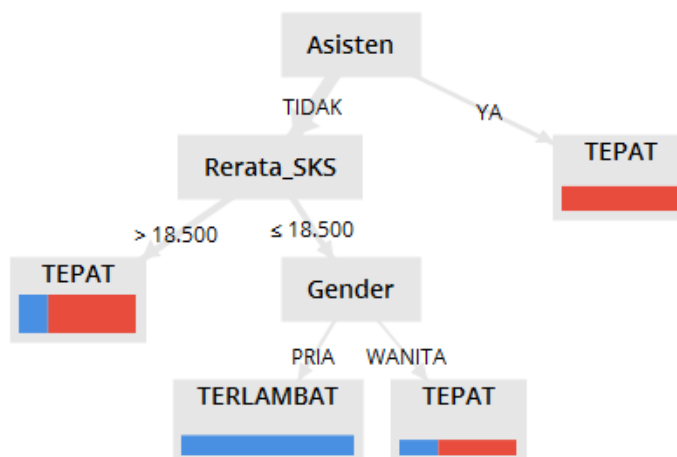
minimal size for split 4 ⓘ

number of prepruning alternati... 3 ⓘ

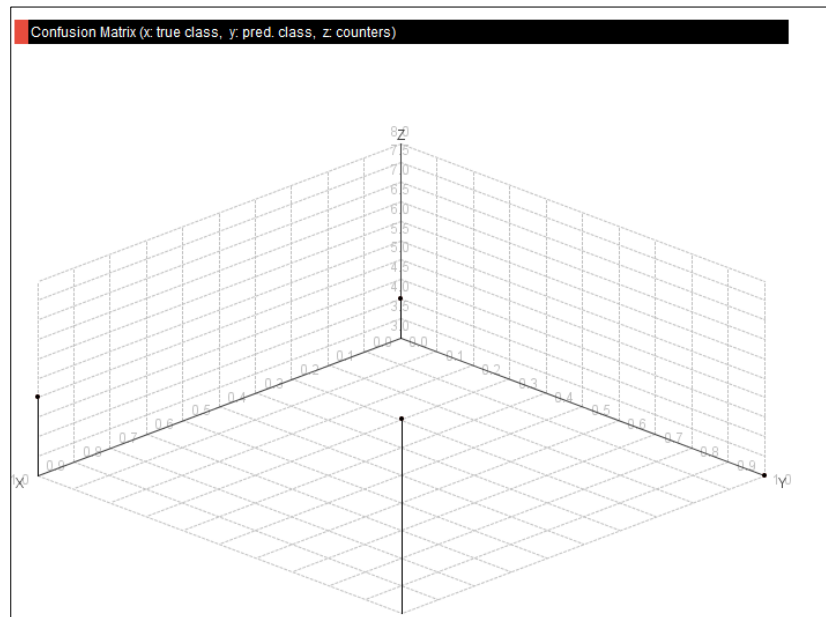
☒ Table View
 ☐ Plot View

accuracy: 60.00% +/- 21.08% (micro average: 60.00%)

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	4	5	44.44%
pred. TEPAT	3	8	72.73%
class recall	57.14%	61.54%	



- b. Cetaklah perspektif plot view dengan model scatter.  
Xaxis = Gender, Yaxis = Asisten, dan Color Column = Lama\_Studi. Nilai Jitter  
bisa diubah-ubah untuk memperoleh pola penyebaran yang lebih jelas.



4. Berdasarkan pohon keputusan dari soal nomor 2, tentukan klasifikasi yang terbentuk berdasarkan kondisinya sesuai dengan simpul-simpulnya.

Klasifikasi yang terbentuk yaitu:

- a. Seorang akan TEPAT(lama\_studi) jika kondisi sebagai berikut:
  - i. Gender = WANITA
  - ii. Gender = PRIA, rerata\_sks > 18 (nilai atribut lain diabaikan)
- b. Seorang akan TERLAMBAT(lama\_studi) jika kondisi sebagai berikut:
  - i. Gender = PRIA, rerata\_sks <= 18 (nilai atribut lain diabaikan)