

Nama : Ihsan Budiono
NIM : L200170119

TUGAS :

1. Berdasarkan pohon keputusan pada percobaan yang kita buat tadi (menggunakan rapid miner). Isikan nilai kelas atribut Bermain_Tenis pada table 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. buka file Jurusan_SMA dengan weka Explorer

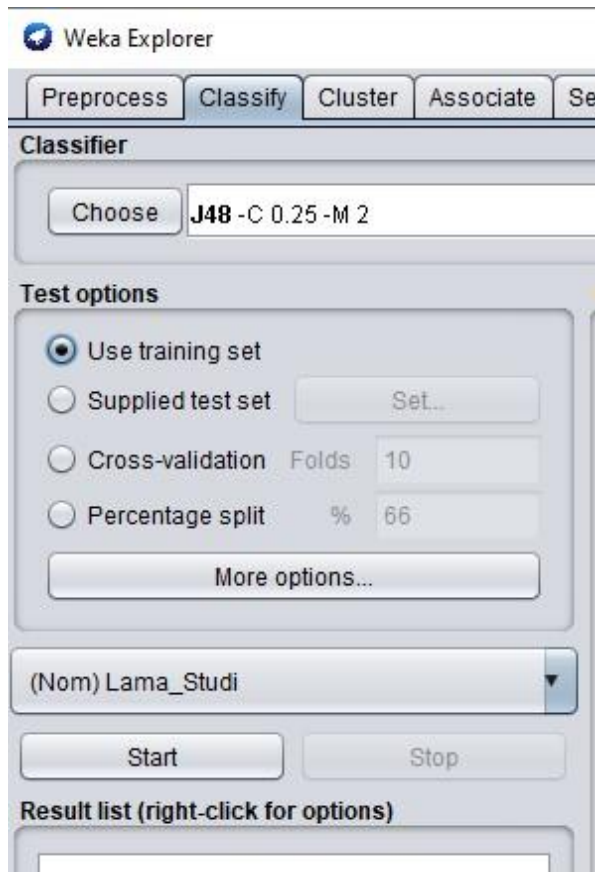
The screenshot shows the Weka Explorer window with the 'Selected attribute' panel active. The panel displays the following information:

- Name: Jurusan_SMA
- Missing: 0 (0%)
- Distinct: 3
- Type: Nominal
- Unique: 0 (0%)

No.	Label	Count	Weight
1	IPS	6	6.0
2	IPA	10	10.0
3	LAIN	4	4.0

Below the table, there is a bar chart visualization showing the distribution of the 'Jurusan_SMA' attribute. The x-axis represents the labels (IPS, IPA, LAIN) and the y-axis represents the count. The bars are colored red and blue, with the red portion representing the majority class (IPA) and the blue portion representing the minority classes (IPS and LAIN).

3. setelah itu buka tab Classify. Tekan tombol choose-Trees-J48. Lalu pilihlah Test Options, pilih Use Training set. Data Jurusan_SMA.arff kita gunakan sebagai data training untuk menciptakan klasifikasi



4. Berdasarkan hasil tersebut dapat diketahui sebagai berikut :

- a) Jumlah simpul daun pada pohon keputusan = 3
- b) Jumlah simpul keseluruhan pada pohon keputusan = 5
- c) Waktu yang dibutuhkan untuk proses pelatihan = 0,02 detik
- d) Tingkat ketepatan klasifikasi = 85%
- e) Tingkat ketidaktepatan klasifikasi = 15%

Classifier output

J48 pruned tree

Gender = PRIA

| Rerata_SKS <= 18: TERLAMBAT (4.0)

| Rerata_SKS > 18: TEPAT (11.0/2.0)

Gender = WANITA: TEPAT (5.0/1.0)

Number of Leaves : 3

Size of the tree : 5

Time taken to build model: 0.02 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	%	
Total Number of Instances	20		

```

Classifier output

=== 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 %
Total Number of Instances          20

=== Detailed Accuracy By Class ===

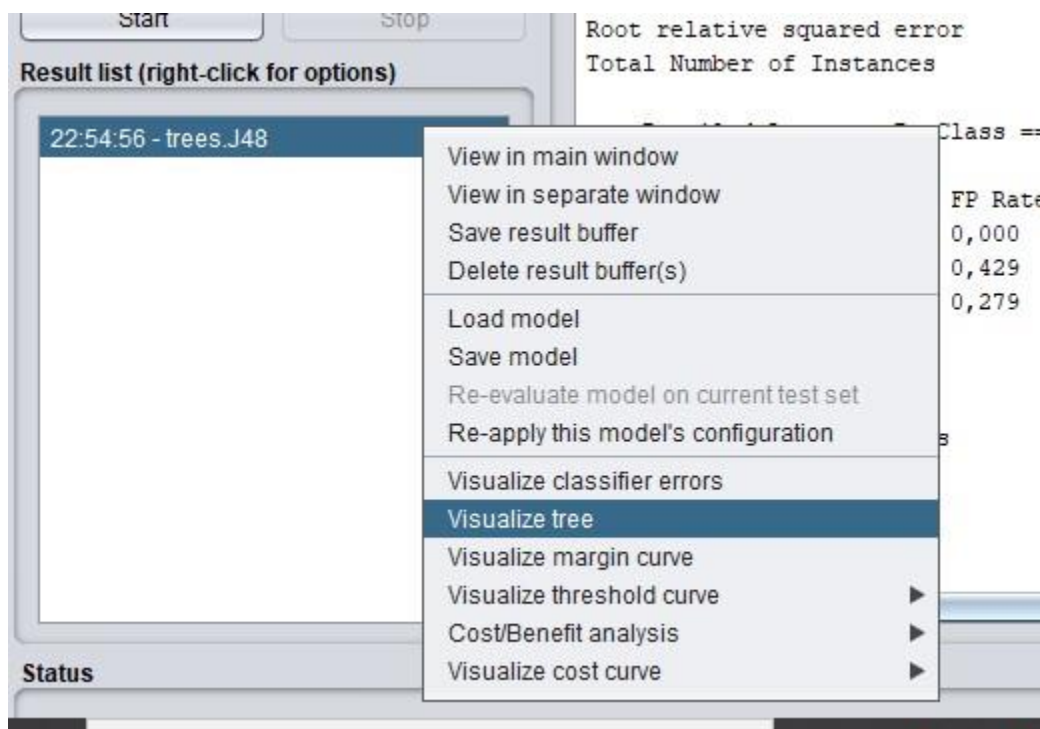
                TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Cla
                0,571    0,000    1,000     0,571    0,727      0,681    0,791    0,751    TER
                1,000    0,429    0,813     1,000    0,897      0,681    0,791    0,816    TEP
Weighted Avg.    0,850    0,279    0,878     0,850    0,837      0,681    0,791    0,793

=== Confusion Matrix ===

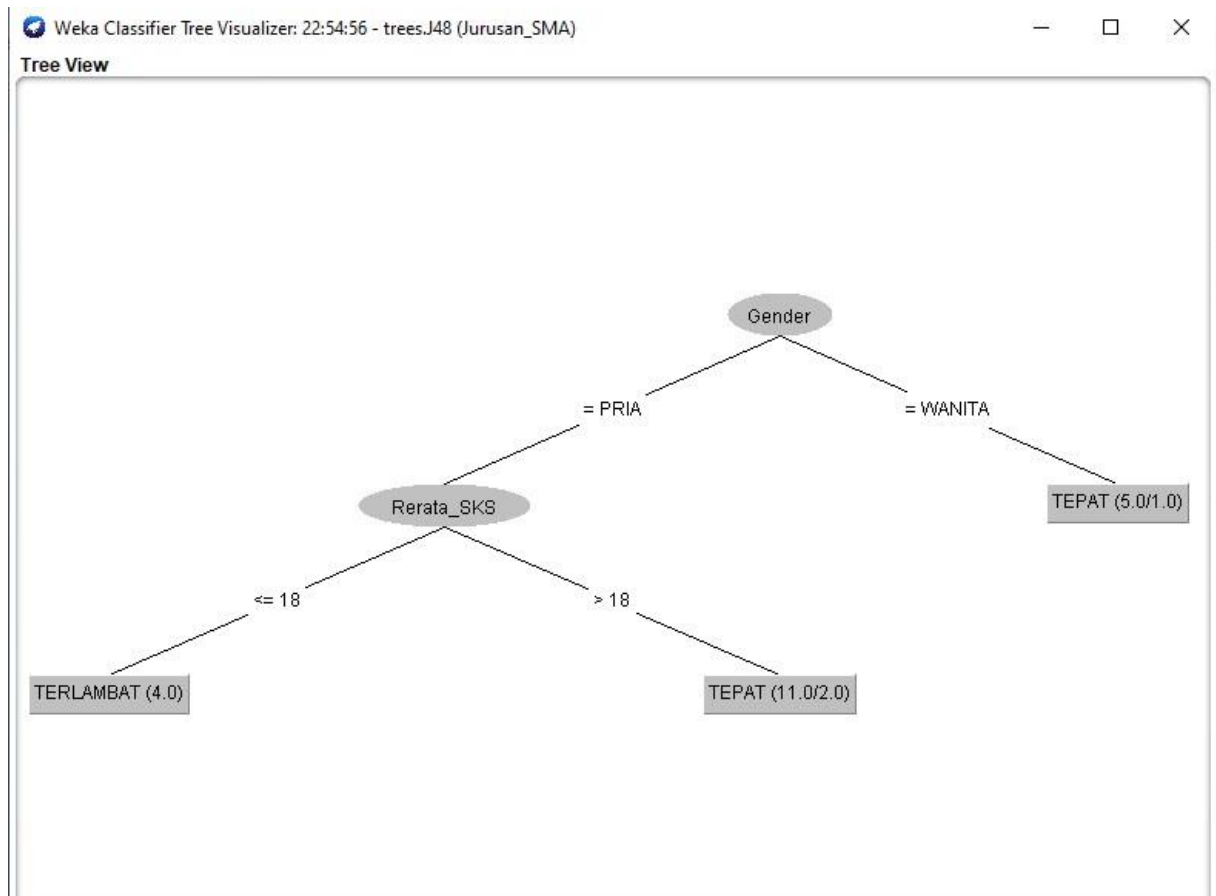
 a  b  <-- classified as
 4  3  |  a = TERLAMBAT
 0 13  |  b = TEPAT

```

5. untuk melihat hasil skema pohon keputusan, kembali ke kolom Result List. Klik kanan pada hasil trees.j480-visualize tree



7. sebuah jendela weka Classifier Tree Visualizer akan ditampilkan. Pada jendela ini akan tampak Tree View, hasil klasifikasi pohon keputusan.



8. berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada, sebagai berikut:

- a) Simpul akar = Gender
- b) Simpul internal = Rerata_SKS
- c) Simpul daun = TERLAMBAT, TEPAT

9. klasifikasi yang terbentuk yaitu :

- a) Seseorang akan Lama Studi (TERLAMBAT) jika kondisi sebagai berikut :
 - i. Gender = PRIA, Rerata_SKS \leq 18 (nilai atribut lain diabaikan)
- b) Seseorang akan Lama Studi (TEPAT) jika kondisi sebagai berikut :
 - i. Gender = PRIA, Rerata_SKS $>$ 18 (nilai atribut lain diabaikan)
 - ii. Gender = WANITA (kondisi lain diabaikan)

Rapid Miner

1. Pada area process view, kita akan mendesain model proses pohon keputusan menggunakan data Mahasiswa_Training. Yang telah kita buat di modul 7

<new process*> - RapidMiner Studio Educational 9.3.001 @ DESKTOP-FQ4H5G4

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Tree (Decision Tree) ExampleSet (//Local Repository/Cuaca_Training)

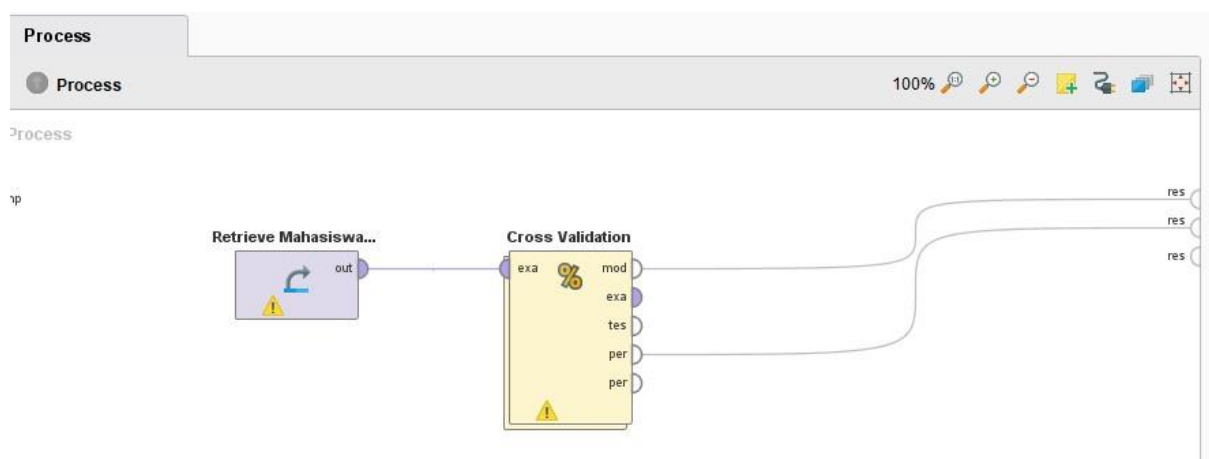
Result History ExampleSet (//Local Repository/Mahasiswa_training) PerformanceVector (Performance)

Open in Turbo Prep Auto Model Filter (20 / 20 examples): all

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	IPS	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	IPA	PRIA	LUAR	17	TIDAK
5	TEPAT	IPA	WANITA	SURAKARTA	17	TIDAK
6	TEPAT	IPA	WANITA	LUAR	18	YA
7	TERLAMBAT	IPA	PRIA	SURAKARTA	18	TIDAK
8	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK
9	TERLAMBAT	IPS	PRIA	LUAR	18	TIDAK
10	TEPAT	LAIN	WANITA	SURAKARTA	18	TIDAK
11	TEPAT	IPA	WANITA	SURAKARTA	19	TIDAK
12	TEPAT	IPS	PRIA	SURAKARTA	20	TIDAK
13	TEPAT	IPS	PRIA	SURAKARTA	19	TIDAK
14	TEPAT	IPA	PRIA	SURAKARTA	19	TIDAK

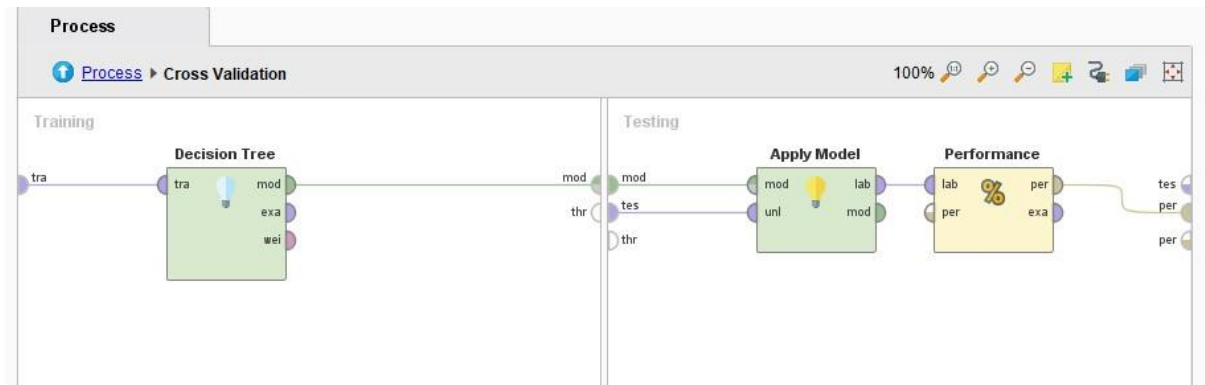
ExampleSet (20 examples. 1 special attribute. 5 regular attributes)

2. Drag Mahasiswa_Training ke area process view, drag pula operator Cross Validation ke area process view. Lalu hubungan tiap portnya sesuai gambar dibawah ini :

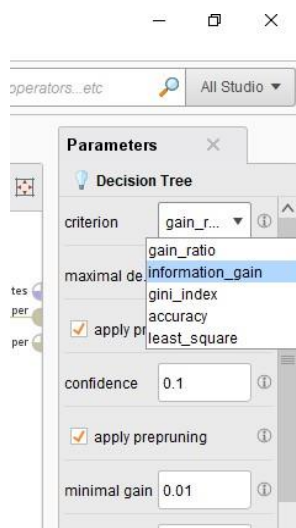


3. klik 2 kali operator Cross validation. Sehingga masuk ke jendela process-Validation. Masukkan operator Decision Tree dalam area Training, operator Apply Model dan

performance dalam area Testing. Hubungkan port input dan output masing- masing operator seperti gambar dibawah ini :



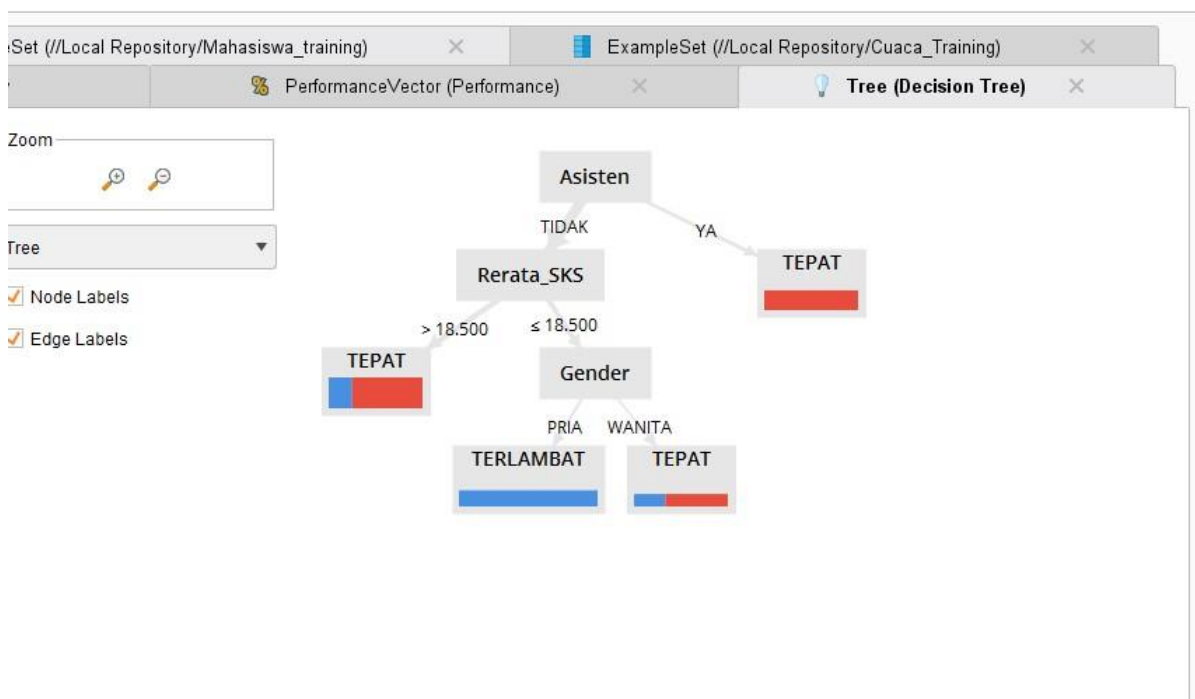
4. klik decision tree pada area training, dan pastikan kriteria (Criterion) yang dipakai adalah information Gain (J48) pada kolom parameter disebalah kanan process view.



5. hasilnya performanceVector

PerformanceVector (Performance)			
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%	

6. hasil Tree (Decision Tree) menunjukkan hasil pohon keputusan dari proses klasifikasi.



7. berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada sebagai berikut :

- Simpul akar = Asisten
- Simpul internal = Rerata_SKS, Gender
- Simpul daun = TEPAT, TERLAMBAT

8. klasifikasi yang terbentuk adalah :

□ Seseorang akan Lama Studi (TEPAT) jika kondisi sebagai berikut :

- i. Asisten = TIDAK, $Rerata_SKS > 18,500$. (nilai atribut lain diabaikan)
- ii. Asisten = TIDAK, $Rerata_SKS \leq 18.500$ Gender(WANITA) = TEPAT
(nilai atribut lain diabaikan) iii.

Asisten = YA.(nilai atribut lain diabaikan)

□ Seseorang akan Lama Studi (TERLAMBAT) jika kondisi sebagai berikut :

- i. Asisten = TIDAK, $Rerata_SKS \leq 18.500$ Gender (PRIA) =
TERLAMBAT (nilai atribut lain diabaikan)

9.

Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
IPA	PRIA	SURAKARTA	19	YA	TEPAT
LAIN	PRIA	SURAKARTA	19	TIDAK	TEPAT
IPS	PRIA	LUAR	17	TIDAK	TERLAMBAT
LAIN	WANITA	SURAKARTA	17	TIDAK	TEPAT
IPA	WANITA	LUAR	18	YA	TEPAT
IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT

10.

