

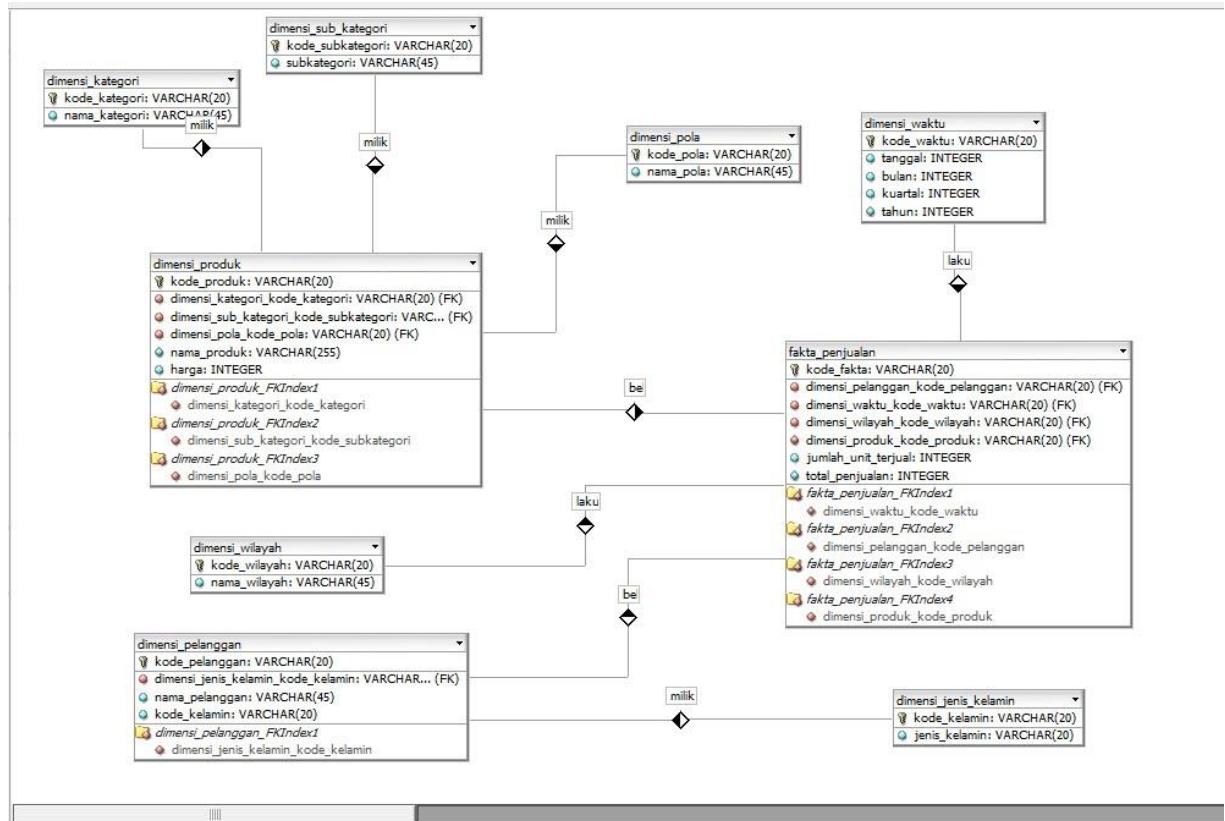
Nama: Yuda Dwi Nurcahyo

Nim : L200170126

LAPORAN PRAKTIKUM DATA WAREHOUSE & DATA MINING

MODUL 1

(Perancangan Star Schema dan Snowflake)



MODUL 5

(Pivot *Table* dan *Chart*)

KEGIATAN

Membuat pivot table kemudian menyimpannya dengan nama ‘Fakta_Penjualan.xls’.

bulan	kuartal	tahun	nama_produk	nama_kategori	nama_subkategori	nama_pola	nama_pelanggan	jenis_kelamin	nama_wilayah	jumlah	harga
2	12	4	2011 Kaos Batik Cap Lukis	Standar	Jarik	Print	Bapak Ketut	PRIA	Bali	2	225000
3	1	1	2012 Batik Standar Cap Tumpal	Batik	Kaos	Cat	Ibu Harini	WANITA	Jawa Timur	14	30000
4	4	2	2012 Celana Standar Cap Warna	Standar	Jarik	Tulis	Ibu Harini	WANITA	Jawa Timur	4	40000
5	4	2	2011 Hem Standar Cap Tumpal	Katun	Hem	Print	Ibu Harini	WANITA	Jawa Timur	3	70000
6	9	3	2012 Bahan Standar Cap Lasem	Standar	Batik	Cap	Bapak Heru	PRIA	Jawa Timur	1	150000
7	5	2	2012 Bahan Standar Cap Garis	Klatun	Hem	Print	Bapak Totok	PRIA	Jawa Timur	3	299000
8	12	4	2011 Bolero Standar Cap Sidomukti	Standar	Bolero	Cap	Ibu Hatamah	WANITA	Jawa Timur	1	225000
9	10	4	2011 Bahan Beludru Cap Makhota	Standar	Sarimbit	Print	Ibu Hatamah	WANITA	Jawa Timur	1	150000
10	1	1	2011 Jarik Standar Tulis Sarimbit	Katun	Kaos	Print	Bapak Imron	PRIA	Jawa Barat	1	60000
11	2	1	2012 Hem Standar Tulis Madura	Standar	Celana	Cap	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
12	3	1	2010 Bahan Lawasan Tulis Tolet	Standar	Celana	Print	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000
13	3	1	2011 Hem Sutra Print Rama	Standar	Bahan	Cap	Ibu Siti Arya	WANITA	Jawa Barat	8	120000
14	12	4	2012 Rok Batik Print Kombinasi	Batik	Rok	Print	Ibu Siti Arya	WANITA	Jawa Barat	1	225000
15	1	1	2012 Jarik Standar Print Sogan	Standar	Jam	Print	Ibu Siti Arya	WANITA	Jawa Barat	44	80000
16	9	3	2012 Celana Standar Print Lasem	Standar	Hem	Cap	Ibu Aini Kasmaji	WANITA	Jawa Tengah	1	100000
17	6	2	2012 Jam Standar Print Lukis	Lawasan	Bahan	Tulis	Ibu Niken	WANITA	Jawa Tengah	1	130000
18	8	3	2011 Sarimbit Standar Print Lukis	Standar	Hem	Tulis	Ibu Atik	WANITA	Jawa Tengah	5	55000
19	4	2	2012 Kaos Katun Print Bola	Standar	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	7	135000
20	6	2	2010 Hem Katun Print Kelengan	Buludru	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	1	50000
21	11	4	2010 Hem Katun Print Kawung	Sutra	Hem	Print	Ibu Tyas	WANITA	Jawa Tengah	5	100000

Hasil pivot table yang menampilkan nama_subkategori, tahun dan jumlah.

Menambahkan tipe summary baru, tambahkan field jumlah kembali ke kotak Value dengan cara klik dan drag, sehingga muncul field baru dengan nama Sum of jumlah2.

	2010	2011	2012	(blank)	Total Sum of jumlah	Total Sum of jumlah2
Sum of jumlah	1	8	8	8	17	17
Sum of jumlah2	1	1	1	1	1	1
					1	1
					1	1
					34	34
					17	17
					44	44
					6	6
					15	15
					1	1
					1	1
	23	23	21	21	93	93
					137	137

Kemudian mengubah nama field sum of jumlah2 menjadi count of jumlah 2 dengan klik sum of jumlah 2 pada PivotTableField dan pilih dialog Value Field Settings.

	2010	2011	2012	(blank)	Total Sum of jumlah	Total Count of jumlah2
Row Labels - Sum of jumlah	1	8	8	2	17	4
Bahan	1	1	1	1	1	1
Batik		1	1		1	1
Bolero					1	1
Celana	17	1	17	1	34	2
Hem	5	1	8	4	17	5
Jam			44	1	44	1
Jarik		2	1	4	6	2
Kaos		1	1	14	15	2
Rok				1	1	1
Sampit		1	1	1	1	1
(blank)						
Grand Total	23	3	21	7	93	10
					137	20

Calculated Field

Caranya klik analyze → pilih fields, items&sets→ calculated field

Pada kotak dialog Insert Calculated Field yang muncul, masukkan nilai berikut kemudian klik tombol OK.

Name : Pendapatan

Formula : = jumlah * harga (Pilih field jumlah kemudian klik Insert Field kemudian ketikkan tanda "*" dan masukkan field harga)

The screenshot shows the Microsoft Excel ribbon with 'PivotTable Tools' selected. A 'PivotTable Fields' pane is open on the right side. In the 'Values' section, 'jumlah' is checked. The main worksheet area shows a PivotTable with columns for 'tahun' (Year) and 'jumlah' (Count of jumlah). The 'Insert Calculated Field' dialog box is open in the center, with 'Name' set to 'Pendapatan' and 'Formula' set to '= jumlah*harga'. The 'Fields' list in the dialog includes 'jumlah', 'harga', and 'jumlah2'.

Field baru, "Sum of Pendapatan" akan muncul pada Pivot Table

		2011			2012				
		Count of jumlah	Sum of Pendapatan	Sum of jumlah	Count of jumlah	Sum of Pendapatan	Sum of jumlah	Count of jumlah	Sum of Pendapatan
6	1	1	500000	8	1	960000	8	2	2120000
7			0			0	1	1	150000
8			0	1	1	225000			0
9	1	1	935000			0	17	1	935000
10	1	1	500000	8	2	4960000	4	2	1586000
11			0			0	44	1	3520000
12			0	2	1	450000	4	1	160000
13			0	1	1	60000	14	1	420000
14			0			0	1	1	225000
15			0	1	1	150000			0
16			0			0			0
11	3	3	15065000	21	7	29400000	93	10	115692000

Operasi Roll Up dan Drill Down

Row Labels	Column Labels		2010		2011			
	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	Count of jumlah2	Sum of Pendapatan	Sum of jumlah	
Bahan	1	1	500000	8	1	960000	0	
Buludru	1	1	500000	0	0	0	0	
Hem Katun Print Kelengan	1	1	500000	0	0	0	0	
Lawasan				0	0	0	0	
Jam Standar Print Lukis				0	0	0	0	
Standar				0	8	1	960000	
Hem Sutra Print Rama				0	8	1	960000	
Kaos Katun Print Bola				0	1	0	0	
Batik				0	1	0	0	
Standar				0	1	1	225000	
Bahan Standar Cap Lasem				0	1	1	225000	
Bolero				0	1	1	225000	
Standar				0	1	1	225000	
Bolero Standar Cap Sidomukti				0	1	1	225000	
Celana	17	1	935000				0	
Standar	17	1	935000				0	
Bahan Lawasan Tulis Toilet	17	1	935000				0	
Hem Standar Tulis Madura			0				0	
Hem	5	1	500000	8	2	4960000		
Katun			0	3	1	210000		
Hem Standar Cap Tumpal			0	3	1	210000		
Klatun			0				0	
Bahan Standar Cap Garis			0				0	
Standar			0	5	1	2750000		
Celana Standar Print Lasem			0	5	1	2750000		
Sarimbhit Standar Print Lukis			0	5	1	2750000		

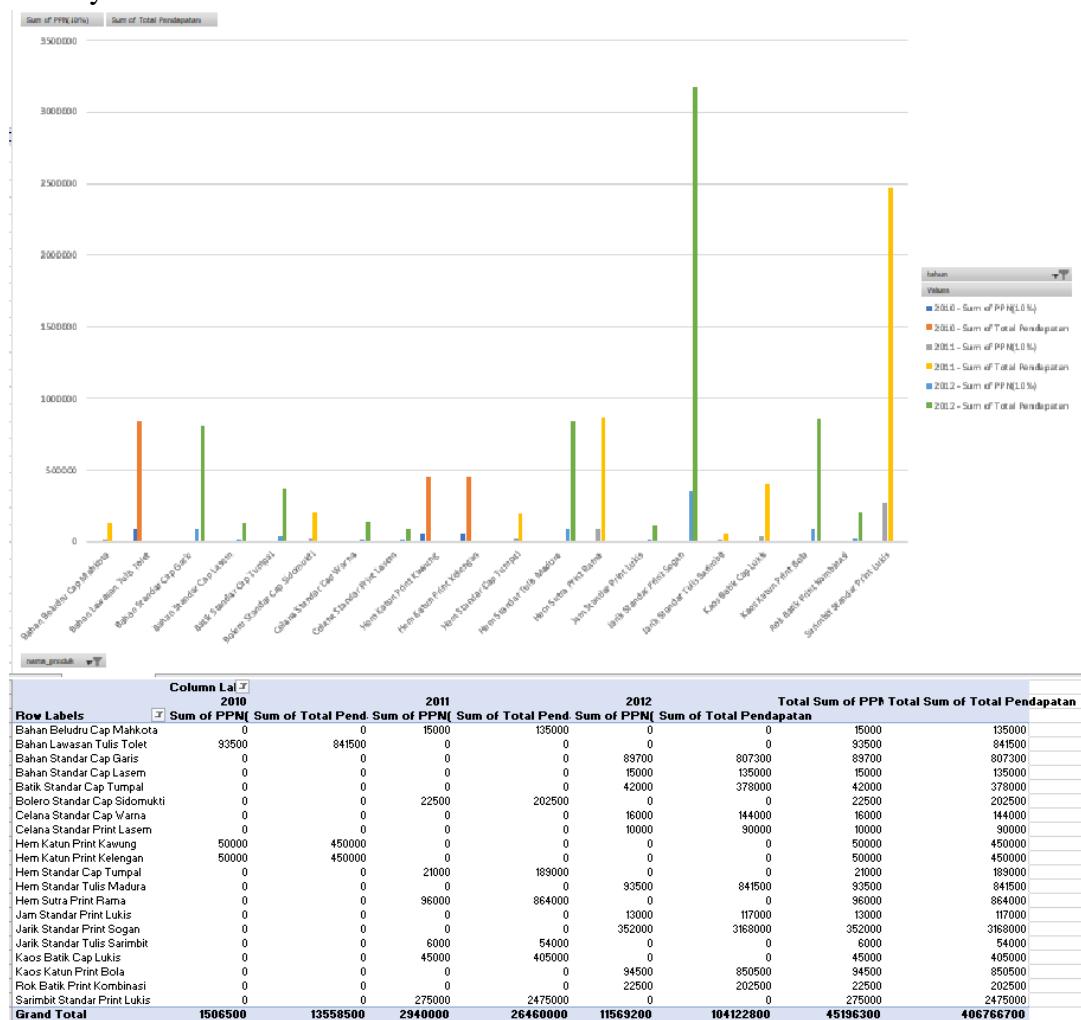
MODUL 5 (Pivot Table dan Chart) TUGAS

Dengan menggunakan PivotTable pada file Fakta_Penjualan.xls tambahkan 2 buah field, yaitu :

- PPN (Pajak Pertambahan Nilai) sebesar 10% dari tiap pendapatan pada Pivot Table.
- Total Penghasilan yang dihitung dari pendapatan dikurangi dengan PPN tersebut.

Sum of Pendapatan	Sum of jumlah	Count of jumlah2	2012		Sum of PPN(10%)	Sum of Total Pendapatan	Total Sum of jumlah	Total Count of jumlah2	Total Sum of Pendapatan	Total Sum of PPN(10%)	Total Sum of Total Pendapatan	
			Sum of Pendapatan	Sum of jumlah								
864000	8	2	2120000	212000	1908000	17	4	15045000	1504500	13540500		
0	1	1	150000	15000	135000	1	1	150000	15000	135000		
202500			0	0	0	0	1	1	225000	22500	202500	
0	17	1	935000	93500	841500	34	2	3740000	374000	3366000		
4464000	4	2	1596000	159600	1436400	17	5	1902300	1902300	17120700		
0	44	1	352000	352000	3168000	44	1	3520000	352000	3168000		
405000	4	1	160000	16000	144000	6	2	1590000	159000	1431000		
54000	14	1	420000	42000	378000	15	2	1350000	135000	1215000		
0	1	1	225000	22500	202500	1	1	225000	22500	202500		
135000			0	0	0	1	1	150000	15000	135000		
26460000	93	10	115692000	11569200	104122800	137	20	451963000	45196300	406766700		

Buatlah PivotTable dan PivotChart untuk melihat PPN dan Total Penghasilan tersebut selama tahun 2010 – 2012. Kategori produk apakah yang memberikan nilai penghasilan terbanyak selama 3 tahun tersebut?



MODUL 6

(Pengenalan *Data Mining*)

	A	B	C	D	E	F	G	H	I
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi		Jumlah Data Jurusan_SMA	
2	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT		6	IPS
3	IPA	PRIA	SURAKARTA	19	YA	TEPAT		10	IPA
4	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT		4	LAIN
5	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT			
6	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT		Jumlah Data Lama_Studi	
7	IPA	WANITA	LUAR	18	YA	TEPAT		13	TEPAT
8	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT		7	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT			
10	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT		Soal C	
11	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT		23	MAX
12	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT		16	MIN
13	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT		18.95	MEAN
14	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT		1.6693838	SDV
15	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT			
16	IPA	PRIA	LUAR	22	YA	TEPAT			
17	LAIN	PRIA	SURAKARTA	16	TIDAK	TERLAMBAT		3	Soal D
18	IPS	PRIA	LUAR	20	TIDAK	TEPAT			
19	LAIN	PRIA	LUAR	23	YA	TEPAT			
20	IPA	PRIA	SURAKARTA	21	YA	TEPAT			
21	IPS	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT			

Menghitung jumlah data yang terdapat dalam kolom Jurusan_SMA

IPS : =COUNTIF(A1:A21,"IPS")

IPA : =COUNTIF(A1:A21,"IPA")

LAIN : =COUNTIF(A1:A21,"LAIN")

Fungsi COUNTIF : menghitung banyaknya variable IPA/IPS/LAIN pada kolom Jurusan_SMA

Hasil :

Jumlah Data Jurusan_SMA	
6	IPS
10	IPA
4	LAIN

Menghitung jumlah data yang terdapat dalam kolom Lama_Studi

TEPAT : =COUNTIF(F1:F21,"TEPAT")

TERLAMBAT : =COUNTIF(F1:F21,"TERLAMBAT")

Fungsi COUNTIF : menghitung banyaknya variable TEPAT/TERLAMBAT pada kolom Lama_Studi

Hasil :

Jumlah Data Lama_Studi	
13	TEPAT
7	TERLAMBAT

Menghitung nilai MAX, MIN, MEAN, Standard Deviation yang terdapat dalam kolom Rerata_SKS

MAX : =MAX(D2:D21)

Fungsi MAX : menghitung nilai maksimum pada kolom Rerata_SKS

MIN : =MIN(D2:D21)

Fungsi MIN : menghitung nilai minimum pada kolom Rerata_SKS

MEAN : =AVERAGE(D2:D21)

Fungsi MEAN : menghitung nilai rata-rata pada kolom Rerata_SKS

Standard Deviation : =STDEV(D2:D21)

Fungsi STDEV : mengukur jumlah variasi atau sebaran sejumlah nilai data pada kolom Rerata_SKS

Hasil :

Soal C	
23	MAX
16	MIN
18.95	MEAN
1.6693838	SDV

Menghitung jumlah data gabungan dari kelas atribut Jurusan_SMA = IPA, Gender = PRIA, Asisten = YA, Lama_Studi = TEPAT.

Formula : =COUNTIFS(A2:A21,"IPA",B2:B21,"PRIA",E2:E21,"YA",F2:F21,"TEPAT")

Fungsi COUNTIFS : menghitung jumlah data atau sel yang memenuhi banyak kriteria

Hasil :

3	Soal D
---	--------

MODUL 7

(Data Preprocessing)

KEGIATAN

Ketentuan variable/atribut :

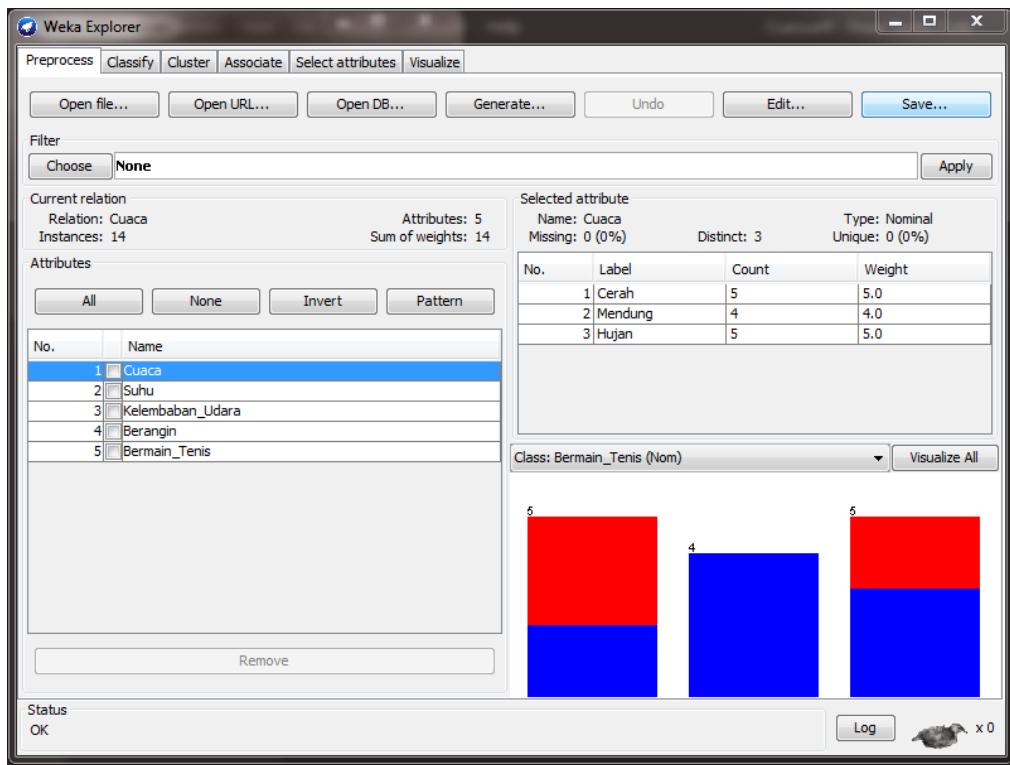
- Variabel Y = Bermain_Tenis
- Variable X1 = Cuaca
- Variable X2 = Suhu
- Variable X3 = Kelembaban_Udara
- Variable X4 = Berangin

Penentuan Data :

- Binomial : jika nilai data hanya ada 2 kemungkinan
- Polynomial : jika nilai data lebih dari 2 kemungkinan, namun tidak terlalu banyak
- Real : jika kemungkinan nilai data sangat beragam dan jumlahnya banyak

```
≡ Cuaca.arff ×

C: > Users > LABSI-05 > Documents > ≡ Cuaca.arff
1  @relation Cuaca
2
3  @attribute Cuaca{Cerah,Mendung,Hujan}
4  @attribute Suhu real
5  @attribute Kelembaban_Udara real
6  @attribute Berangin{YA,TIDAK}
7  @attribute Bermain_Tenis {YA,TIDAK}
8
9  @data
10 Cerah,85,85,TIDAK,TIDAK
11 Cerah,80,90,YA,TIDAK
12 Mendung,83,86,TIDAK,YA
13 Hujan,70,96,TIDAK,YA
14 Hujan,68,80,TIDAK,YA
15 Hujan,65,70,YA,TIDAK
16 Mendung,64,65,YA,YA
17 Cerah,72,95,TIDAK,TIDAK
18 Cerah,69,70,TIDAK,YA
19 Hujan,75,80,TIDAK,YA
20 Cerah,75,70,YA,YA
21 Mendung,72,90,YA,YA
22 Mendung,81,75,TIDAK,YA
23 Hujan,71,91,YA,TIDAK
24
```

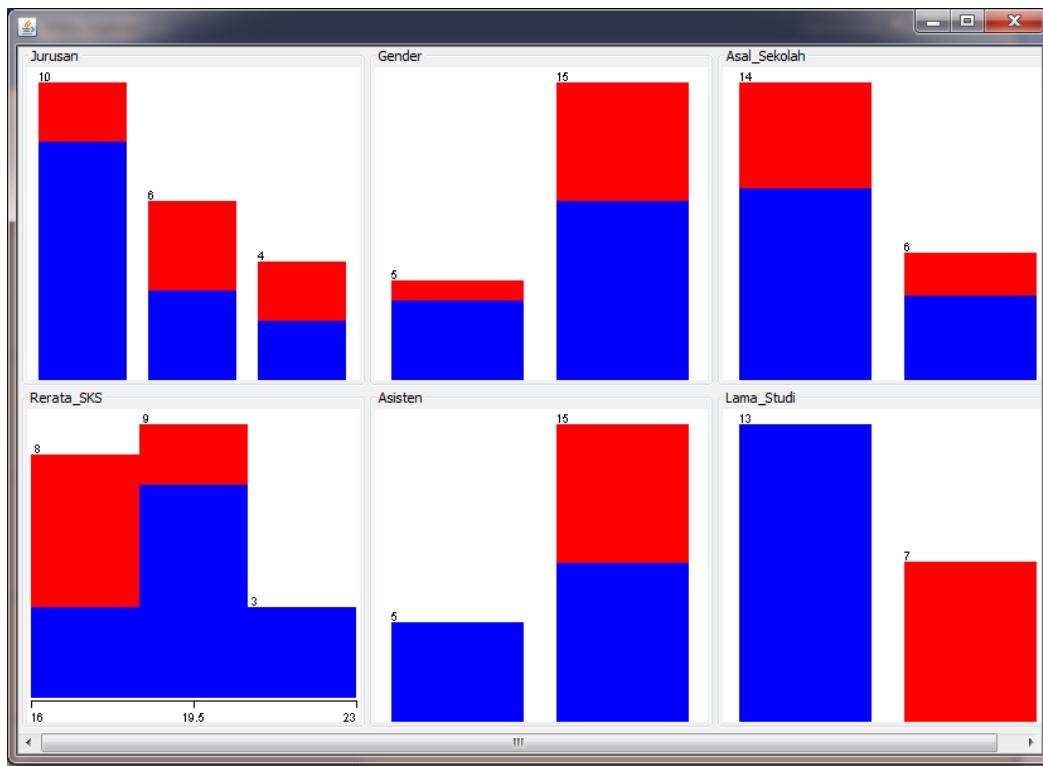
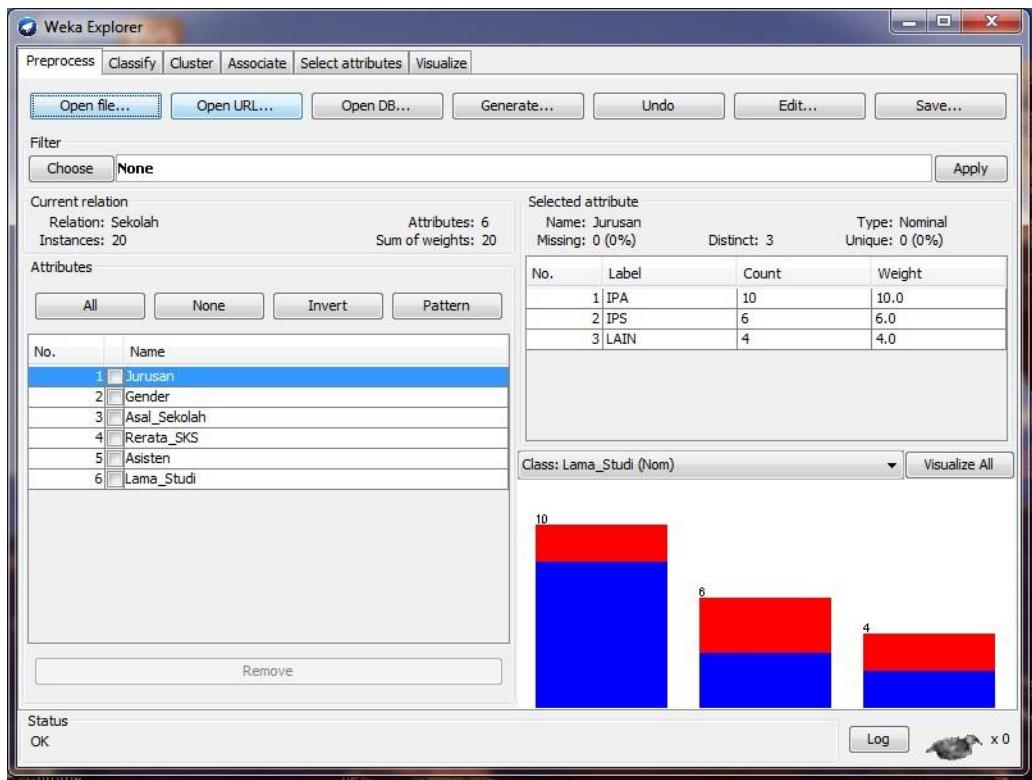


MODUL 7

(Data Preprocessing)

TUGAS

```
≡ Sekolah.arff ×  ≡ Cuaca.arff ×  
C: > Users > LABSI-05 > Documents > ≡ Sekolah.arff  
1  @relation Sekolah  
2  
3  @attribute Jurusan {IPA, IPS, LAIN}  
4  @attribute Gender {WANITA, PRIA}  
5  @attribute Asal_Sekolah {SURAKARTA, LUAR}  
6  @attribute Rerata_SKS real  
7  @attribute Asisten {YA, TIDAK}  
8  @attribute Lama_Studi {TEPAT, TERLAMBAT}  
9  
10 @data  
11 IPA,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT  
12 IPA,PRIA,SURAKARTA,19,YA,TEPAT  
13 LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT  
14 IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT  
15 IPA,WANITA,SURAKARTA,17,TIDAK,TEPAT  
16 IPA,WANITA,LUAR,18,YA,TEPAT  
17 IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT  
18 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT  
19 IPS,PRIA,LUAR,18,TIDAK,TERLAMBAT  
20 LAIN,WANITA,SURAKARTA,18,TIDAK,TEPAT  
21 IPA,WANITA,SURAKARTA,19,TIDAK,TEPAT  
22 IPS,PRIA,SURAKARTA,20,TIDAK,TEPAT  
23 IPS,PRIA,SURAKARTA,19,TIDAK,TEPAT  
24 IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT  
25 IPA,PRIA,LUAR,22,YA,TEPAT  
26 LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT  
27 IPS,PRIA,LUAR,20,TIDAK,TEPAT  
28 LAIN,PRIA,LUAR,23,YA,TEPAT  
29 IPA,PRIA,SURAKARTA,21,YA,TEPAT  
30 IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT  
31
```



Jumlah atribut bertipe binomial yaitu 4

- Gender {WANITA, PRIA}
- Asal_Sekolah {SURAKARTA, LUAR}
- Asisten {YA, TIDAK}
- Lama_Studi {TEPAT, TERLAMBAT}

Jumlah atribut bertipe polynomial yaitu 1

- Jurusan {IPA, IPS, LAIN}

Jumlah atribut bertipe real yaitu 1

- Rerata_SKS real

Pada atribut Rerata_SKS, besarnya nilai :

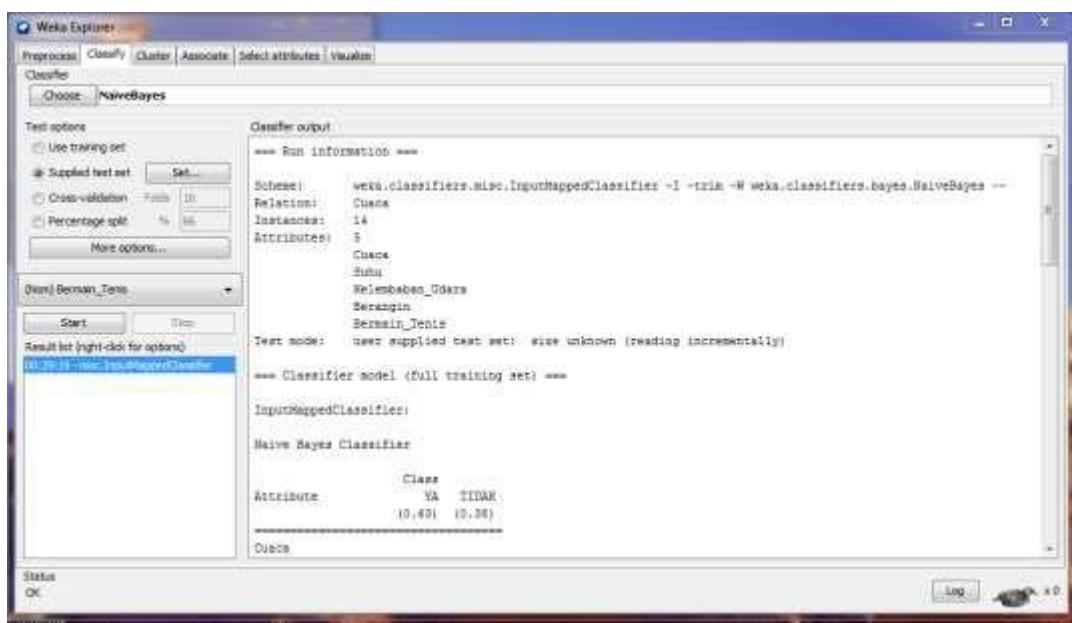
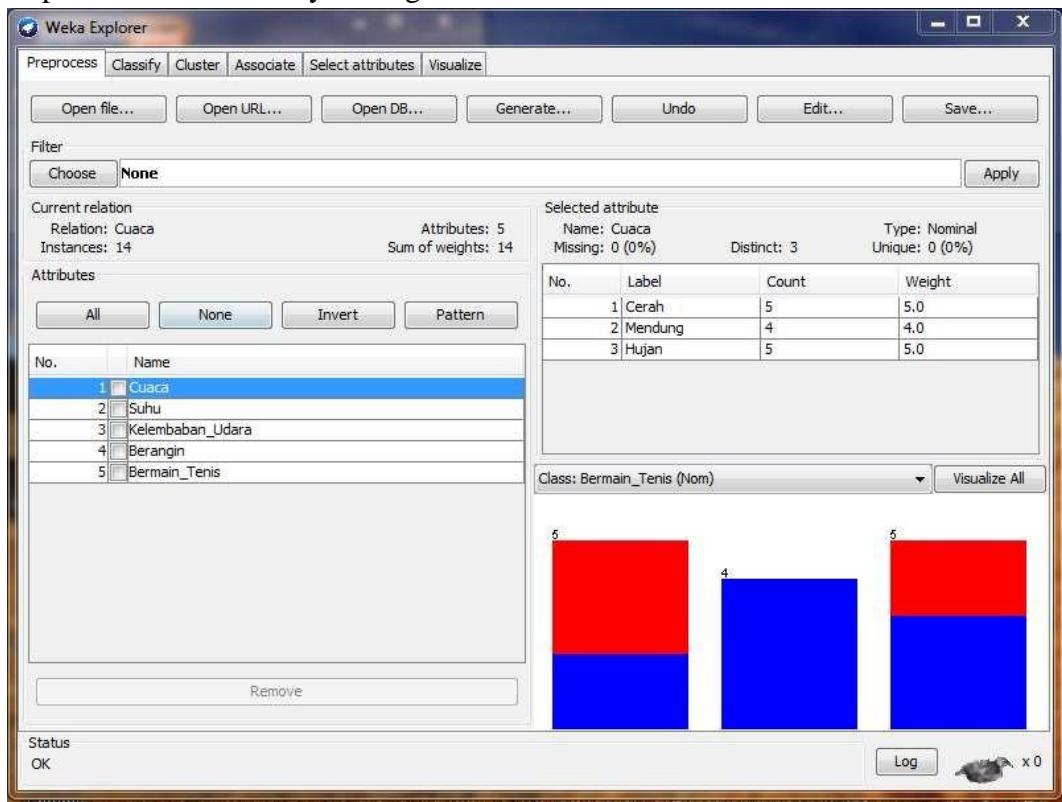
1. Maximum : 23
2. Minimum : 16
3. Mean : 18.95
4. StdDev : 1.6693838

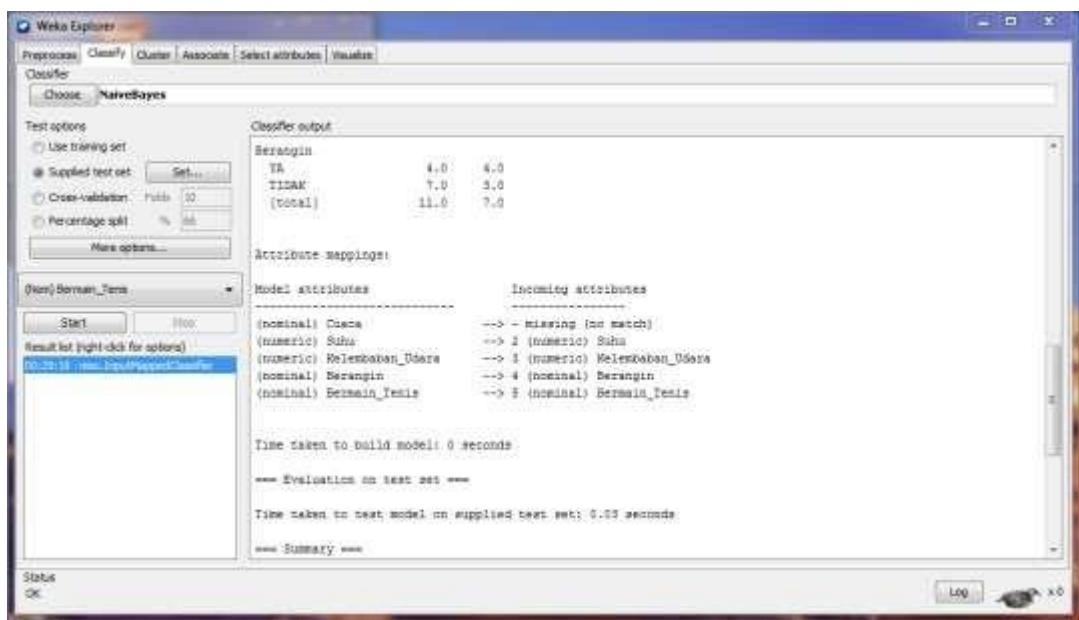
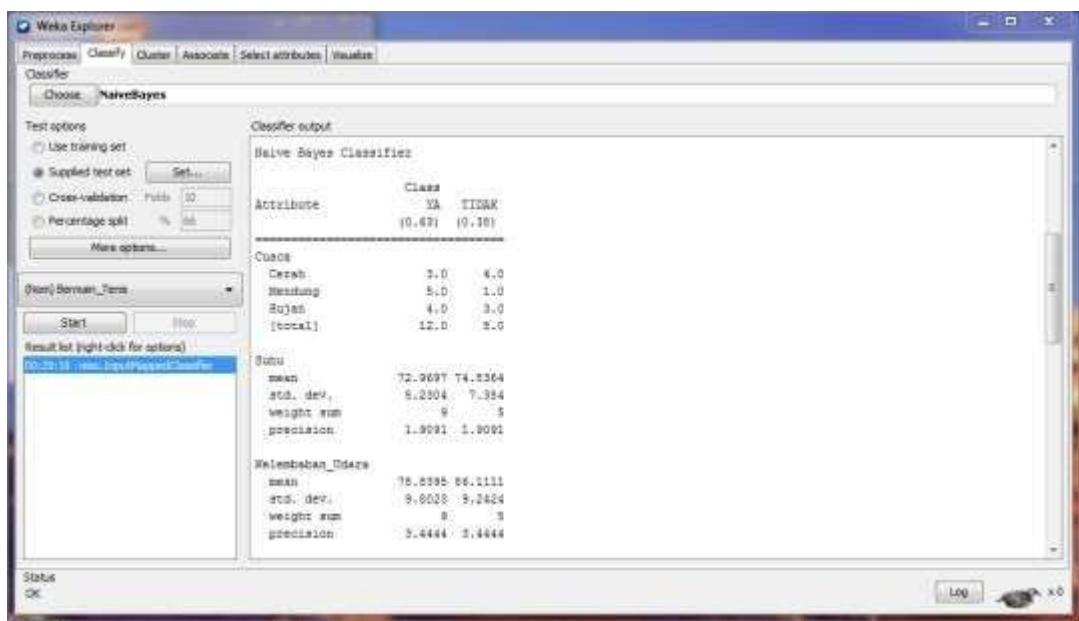
MODUL 8

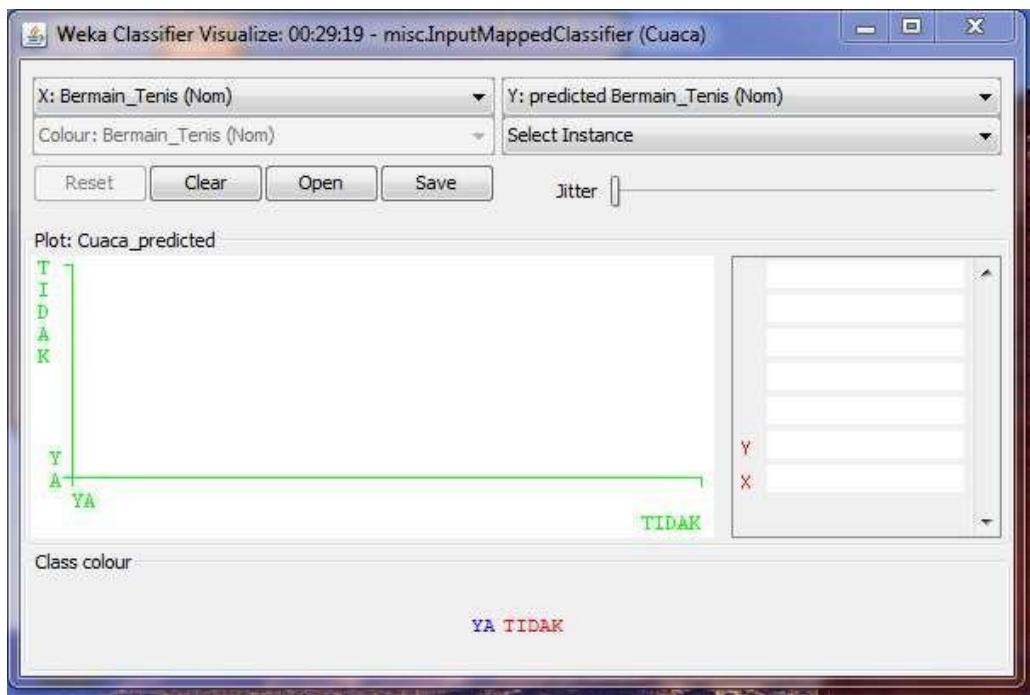
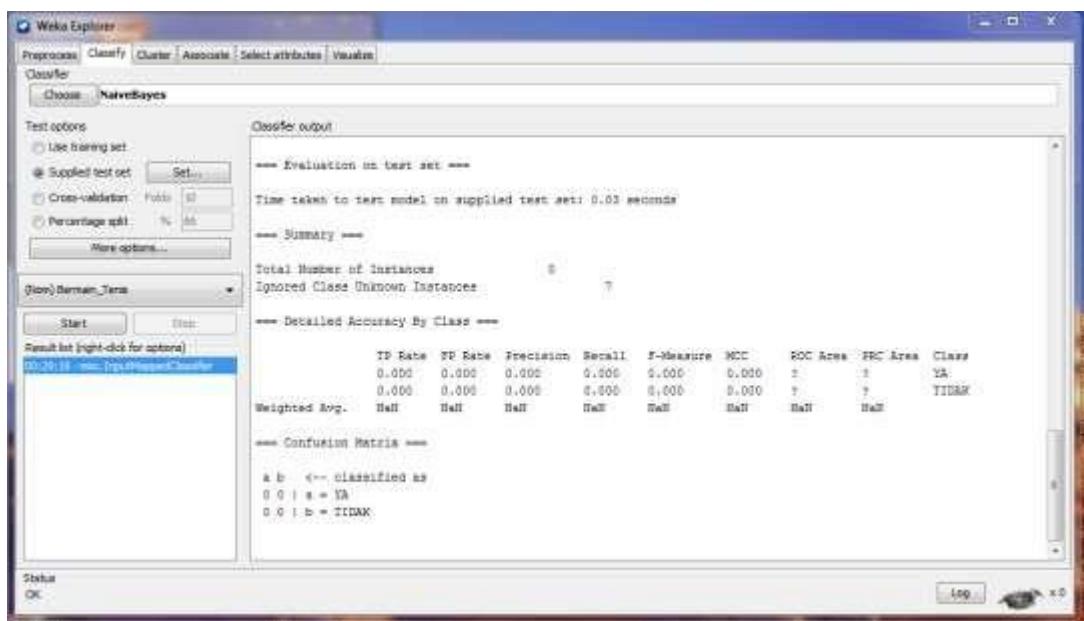
(Klasifikasi : Naïve Bayes)

KEGIATAN

Implementasi Naïve Bayes dengan Weka







ARFF-Viewer - C:\Users\LABSI-05\Documents\HasilPrediksi.arff

File Edit View
HasilPrediksi.arff

Relation: Cuaca_predicted

No.	1: Cuaca Nominal	2: Suhu Numeric	3: Kelembaban_Udara Numeric	4: Berangin Nominal	5: prediction margin Numeric	6: predicted Bermain_Tenis Nominal	7: Bermain_Tenis Nominal
1	Cerah	75.0	65.0	TIDAK	0.762765	YA	
2	Cerah	80.0	68.0	YA	0.087878	YA	
3	Cerah	83.0	87.0	YA	-0.676866	TIDAK	
4	Mendung	70.0	96.0	TIDAK	0.628523	YA	
5	Mendung	68.0	81.0	TIDAK	0.833996	YA	
6	Hujan	65.0	75.0	YA	0.253733	YA	
7	Hujan	64.0	85.0	YA	-0.160143	TIDAK	

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Training ▾ Cell range: A:E Select All Define header row: 1

	A	B	C	D	E
1	Cuaca	Sunu	Kelembaban_udara	Berangin	Bermain_tenis
2	Cerah	85.000	85.000	TIDAK	TIDAK
3	Cerah	80.000	90.000	YA	TIDAK
4	Mendung	83.000	86.000	TIDAK	YA
5	Hujan	70.000	96.000	TIDAK	YA
6	Hujan	68.000	80.000	TIDAK	YA
7	Hujan	65.000	70.000	YA	TIDAK
8	Mendung	64.000	65.000	YA	YA
9	Cerah	72.000	95.000	TIDAK	TIDAK
10	Cerah	69.000	70.000	TIDAK	YA
11	Hujan	75.000	80.000	TIDAK	YA
12	Cerah	75.000	70.000	YA	YA
13	Mendung	72.000	90.000	YA	YA
14	Mendung	81.000	75.000	TIDAK	YA
15	Hujan	71.000	91.000	YA	TIDAK

← Previous → Next ✖ Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Cuaca polynomial	Suhu integer	Kelembaban_u... integer	Berangin polynomial	Bermain_Tenis polynomial
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	YA
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

 no problems.

 [Previous](#)  [Next](#)  [Cancel](#)

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Cuaca polynominal	Suhu integer	Kelembaban_u... integer	Berangin polynominal	Bermain_Tenis <i>binomial</i> label
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Mendung	83	86	TIDAK	YA
4	Hujan	70	96	TIDAK	YA
5	Hujan	68	80	TIDAK	YA
6	Hujan	65	70	YA	TIDAK
7	Mendung	64	65	YA	YA
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

✓ no problems.

← Previous Next → Cancel

New project - RapidMiner Studio Trial 3.8.0 (Windows-25-PC)

File Edit Execute View Connectors Settings Extensions Help

Import Data - Format your columns.

Result History ExampleSet / Local Repository/DataCuaca_Training

Repository ImportData

Data

Row No.	Bermain_Tenis	Cuaca	Suhu	Kelembaban_u...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	81	YA

Completed (14 examples, 1 special attribute, 4 regular attributes)

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Testing ▾ Cell range: A:D Select All Define header row: 1

A	B	C	D
1 Cuaca	Suhu	Kelembaban_udara	Berangin
2 Cerah	75.000	65.000	TIDAK
3 Cerah	80.000	68.000	YA
4 Cerah	83.000	87.000	YA
5 Mendung	70.000	96.000	TIDAK
6 Mendung	68.000	81.000	TIDAK
7 Hujan	65.000	75.000	YA
8 Hujan	64.000	85.000	YA

← Previous → Next ✖ Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

Cuaca polynominal	Suhu integer	Kelembaban_udara integer	Berangin polynominal
1 Cerah	75	65	TIDAK
2 Cerah	80	68	YA
3 Cerah	83	87	YA
4 Mendung	70	96	TIDAK
5 Mendung	68	81	TIDAK
6 Hujan	65	75	YA
7 Hujan	64	85	YA

✓ no problems

← Previous → Next ✖ Cancel

File Edit Projects View Connections Settings Extensions Help

Result History ExampleSet (Local Repository>DataCusca_Testing) ExampleSet (Local Repository>DataCusca_Training)

Data Statistics Visualizations Databases

Open H Tunes Prep Add Model

Row No. Cusca Sales Retention... Demographic

1 Corp 75 66 YDKK

2 Corp 80 66 YK

3 Corp 83 81 YA

4 Mngng 70 66 YDKK

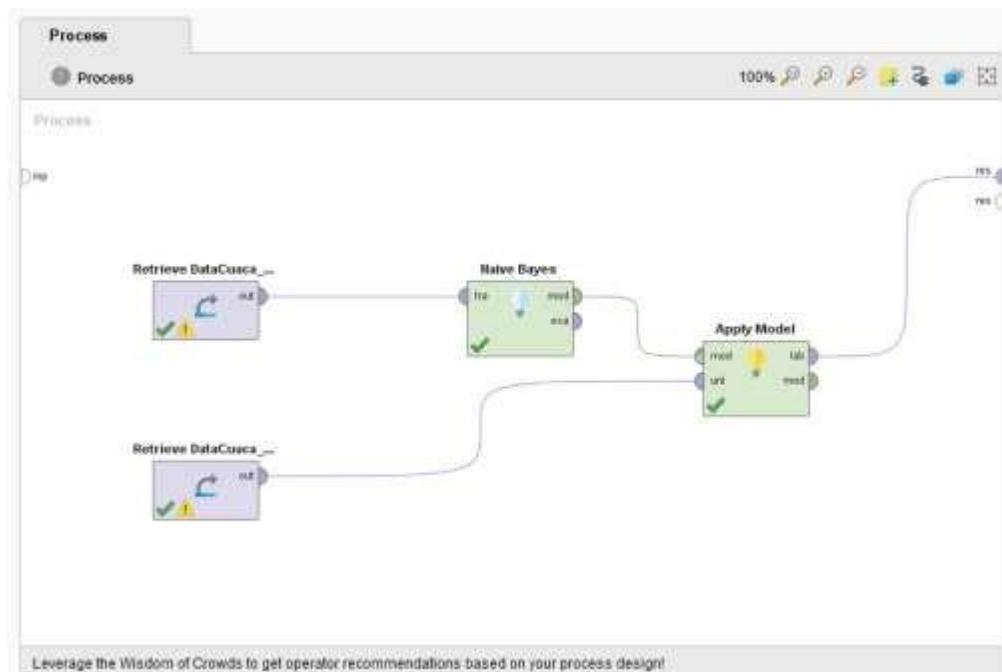
5 Mngng 88 81 YDKK

6 Human 95 75 YA

7 Human 54 66 YK

Completed (7 examples, 0 special attributes, 4 regular attributes)

Repository ImportData MR Training Resources Scripts Community Samples DB Local Repository Connectors DataCusca_Testing DataCusca_Training



Result History ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)

Open in: Turbo Prep Auto Model Filter (7 / 7 examples) all

Data Statistics Visualizations Annotations

Row No.	prediction(Bernmain_Tenis)	confidence(TIDAK)	confidence(YA)	Cuaca	Suhu	Kelembaban_udara	Berangin
1	YA	0.154	0.846	Cerah	75	85	TIDAK
2	YA	0.498	0.502	Cerah	80	88	YA
3	TIDAK	0.856	0.144	Cerah	83	87	YA
4	YA	0.019	0.981	Mendung	70	96	TIDAK
5	YA	0.007	0.993	Mendung	68	81	TIDAK
6	YA	0.371	0.629	Hujan	66	75	YA
7	TIDAK	0.588	0.432	Hujan	64	85	YA

Result History ExampleSet (Apply Model)

Name Type Missing Statistics Filter (7/7 attributes) Search by Attribute

Data Statistics Visualizations Annotations

prediction(Bernmain_Tenis)	Binomial	0	Min: TIDAK (2) Max: YA (5) Average: YA (5), TIDAK (2)
confidence(TIDAK)	Real	0	Min: 0.007 Max: 0.556 Average: 0.353
confidence(YA)	Real	0	Min: 0.144 Max: 0.993 Average: 0.647
Cuaca	Polynomial	0	Min: Mendung (2) Max: Cerah (3) Values: Cerah (3), Hujan (2), Mendung (2)
Suhu	Integer	0	Min: 64 Max: 83 Average: 72.143
Kelembaban_udara	Integer	0	Min: 65 Max: 96 Average: 79.571
Berangin	Polynomial	0	Min: TIDAK (3) Max: YA (4) Values: YA (4), TIDAK (3)

Browsing attributes 1-7 Examples: 7 Special Attributes: 3 Regular Attributes: 4

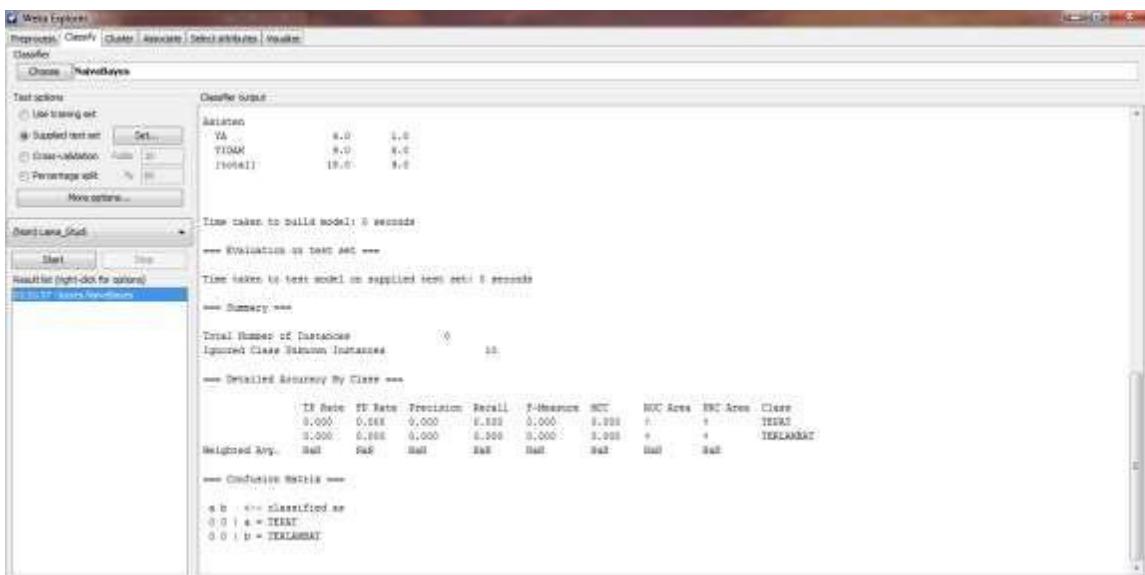
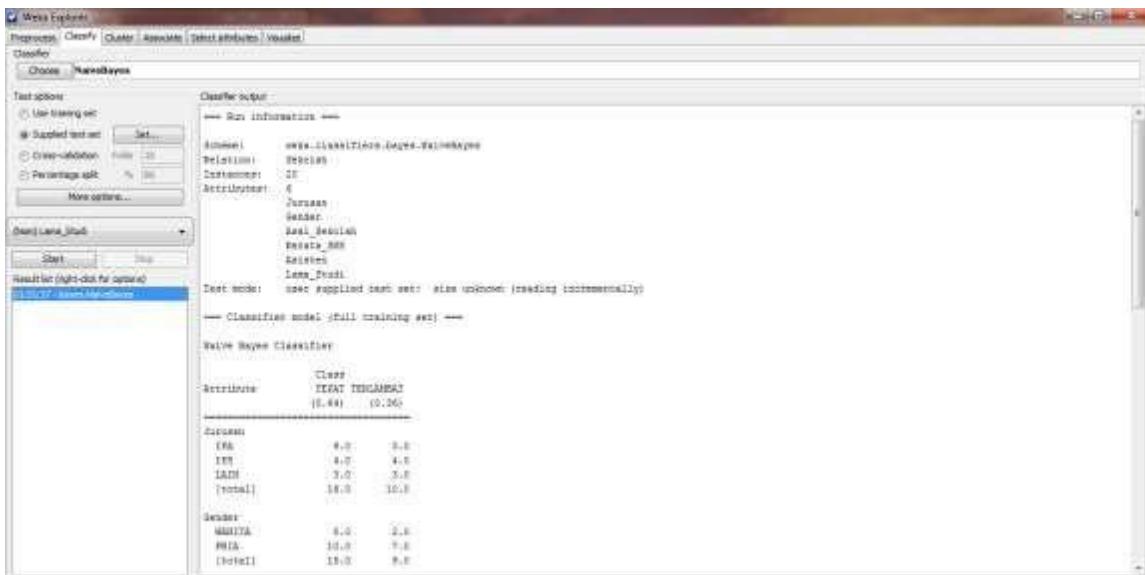
MODUL 8
(Klasifikasi : Naïve Bayes)
TUGAS

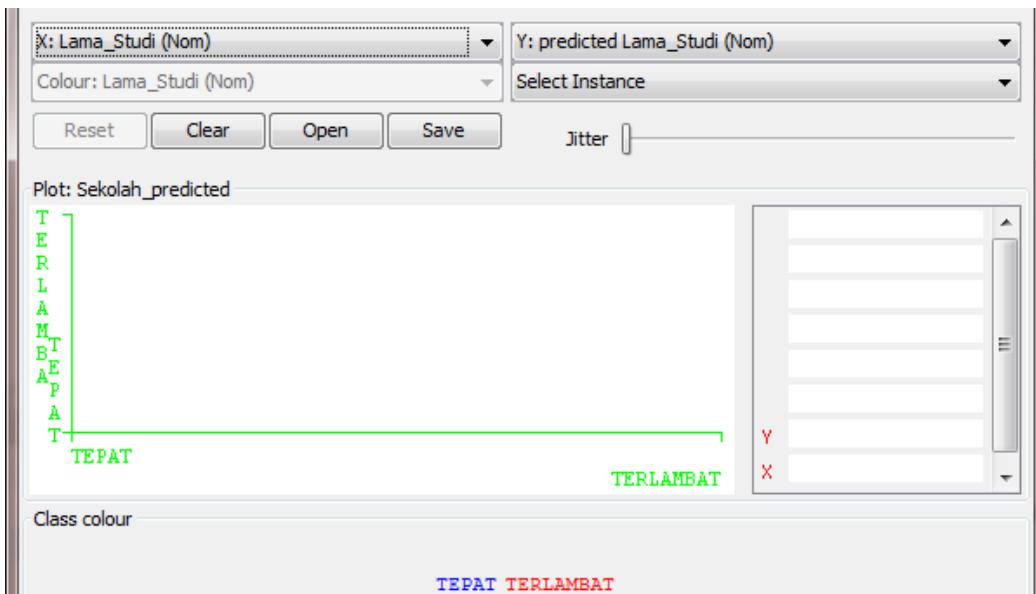
Data Training

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT
15	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
16	IPA	PRIA	LUAR	22	YA	TEPAT
17	LAIN	PRIA	SURAKARTA	16	TIDAK	TERLAMBAT
18	IPS	PRIA	LUAR	20	TIDAK	TEPAT
19	LAIN	PRIA	LUAR	23	YA	TEPAT
20	IPA	PRIA	SURAKARTA	21	YA	TEPAT
21	IPS	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT

Data Testing

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	LAIN	WANITA	SURAKARTA	18	TIDAK	?
3	IPA	PRIA	SURAKARTA	19	YA	?
4	LAIN	PRIA	SURAKARTA	19	TIDAK	?
5	IPS	PRIA	LUAR	17	TIDAK	?
6	LAIN	WANITA	SURAKARTA	17	TIDAK	?
7	IPA	WANITA	LUAR	18	YA	?
8	IPA	PRIA	SURAKARTA	18	TIDAK	?
9	IPA	PRIA	SURAKARTA	19	TIDAK	?
10	IPS	PRIA	LUAR	18	TIDAK	?
11	LAIN	WANITA	SURAKARTA	18	TIDAK	?
12						





ARFF-Viewer - C:\Users\LABSI-05\Documents\HasilTugas.arff

File Edit View

HasilTugas.arff

Relation: Sekolah_predicted

No.	1: Jurusan Nominal	2: Gender Nominal	3: Asal_Sekolah Nominal	4: Rerata_SKS Numeric	5: Asisten Nominal	6: prediction margin Numeric	7: predicted Lama_Studi Nominal	8: Lama_Studi Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	
2	IPA	PRIA	SURAKARTA	19.0	YA	0.836469	TEPAT	
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK	-0.175169	TERLAMBAT	
4	IPS	PRIA	LUAR	17.0	TIDAK	-0.713206	TERLAMBAT	
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK	-0.546846	TERLAMBAT	
6	IPA	WANITA	LUAR	18.0	YA	0.757815	TEPAT	
7	IPA	PRIA	SURAKARTA	18.0	TIDAK	-0.125076	TERLAMBAT	
8	IPA	PRIA	SURAKARTA	19.0	TIDAK	0.356012	TEPAT	
9	IPS	PRIA	LUAR	18.0	TIDAK	-0.588286	TERLAMBAT	
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK	-0.375862	TERLAMBAT	

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Jurusan_SMA * <small>polynominal</small>	Gender * <small>polynominal</small>	Asal_Sekolah * <small>polynominal</small>	Rerata_SKS * <small>integer</small>	Asisten * <small>polynominal</small>	Lama_Studi * <small>binominal label</small>
1	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
2	IPA	PRIA	SURAKARTA	19	YA	TEPAT
3	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
4	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
5	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
6	IPA	WANITA	LUAR	18	YA	TEPAT
7	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
8	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
9	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
10	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
11	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
12	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT

✓ no problems.

← Previous → Next X Cancel

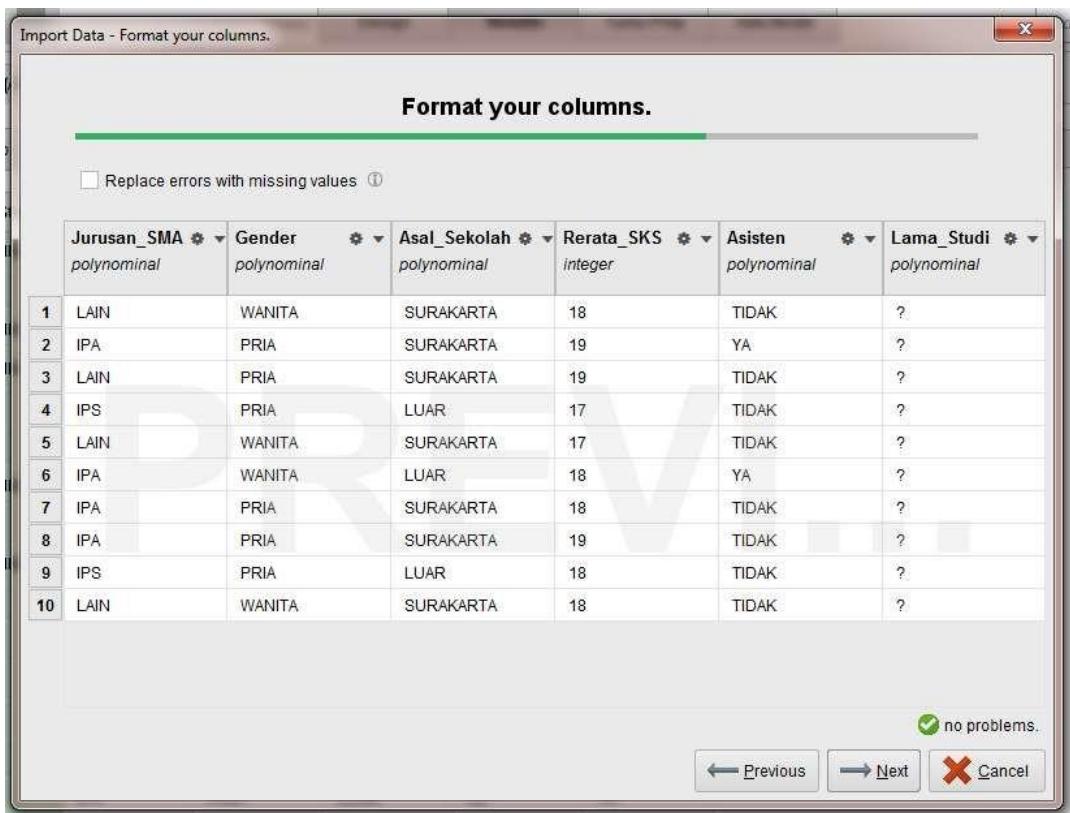
Result History

ExampleSet (Local Repository/DataSekolah_Training)

Data Open in Turbo Prep Auto Model

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	18	TIDAK
15	TEPAT	IPA	PRIA	LUAR	22	YA

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



Result History

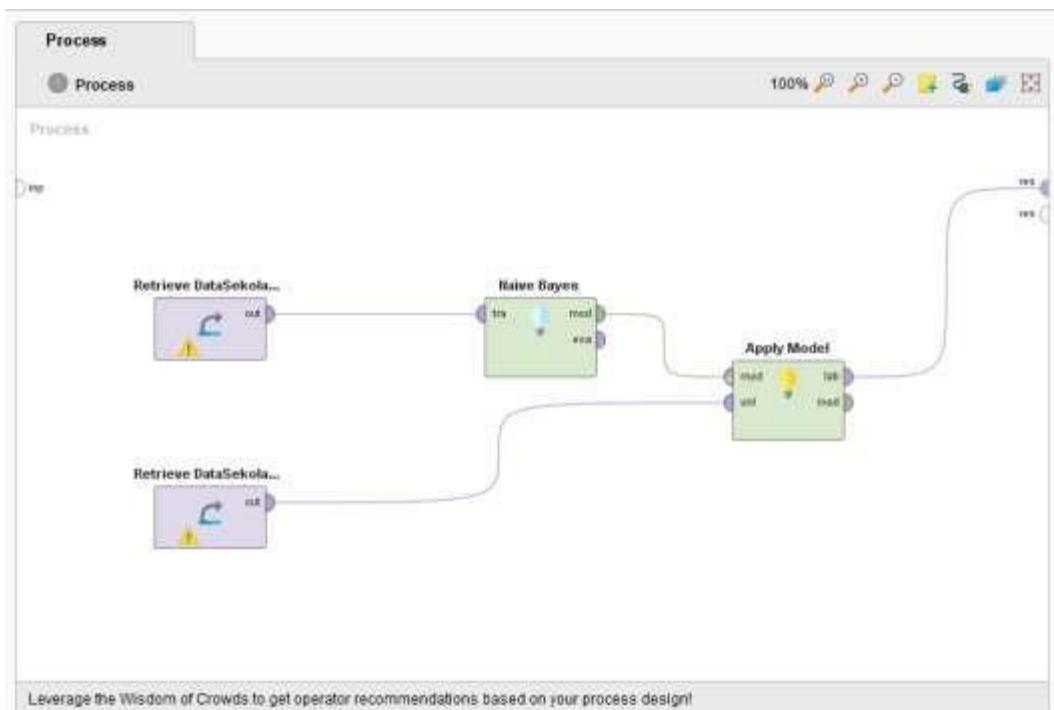
ExampleSet (iLocal Repository/DataSekolah_Testing) ExampleSet (iLocal Repository/DataSekolah_Training)

Open in  

Filter (10 / 10 examples): aR

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
1	LAIN	WANITA	SURAKARTA	18	TIDAK	?
2	IPA	PRIA	SURAKARTA	19	YA	?
3	LAIN	PRIA	SURAKARTA	19	TIDAK	?
4	IPS	PRIA	LUAR	17	TIDAK	?
5	LAIN	WANITA	SURAKARTA	17	TIDAK	?
6	IPA	WANITA	LUAR	18	YA	?
7	IPA	PRIA	SURAKARTA	18	TIDAK	?
8	IPA	PRIA	SURAKARTA	19	TIDAK	?
9	IPS	PRIA	LUAR	18	TIDAK	?
10	LAIN	WANITA	SURAKARTA	18	TIDAK	?

ExampleSet (10 examples, 0 special attributes, 6 regular attributes)



Open in [Turbo Prep](#) [Auto Model](#)

Filter (10 / 10 examples): [all](#)

Row No.	prediction L...	confidence ...	confidence ...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
1	TERLAMBAT	0.648	0.352	LANJ	WANITA	SURAKARTA	18	TIDAK	?
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	18	YA	?
3	TERLAMBAT	0.650	0.350	LANJ	PRIA	SURAKARTA	18	TIDAK	?
4	TERLAMBAT	0.866	0.132	IPS	PRIA	LUAR	17	TIDAK	?
5	TERLAMBAT	0.730	0.262	LANJ	WANITA	SURAKARTA	17	TIDAK	?
6	TEPAT	0.005	0.995	IPA	WANITA	LUAR	18	YA	?
7	TERLAMBAT	0.547	0.453	IPA	PRIA	SURAKARTA	18	TIDAK	?
8	TEPAT	0.321	0.679	IPA	PRIA	SURAKARTA	19	TIDAK	?
9	TERLAMBAT	0.611	0.189	IPS	PRIA	LUAR	18	TIDAK	?
10	TERLAMBAT	0.648	0.352	LANJ	WANITA	SURAKARTA	18	TIDAK	?

Rerata confidence Lama_studi dengan nilai TEPAT

confidence(...
0.848
0.005
0.850
0.868
0.738
0.005
0.547
0.321
0.811
0.648

$$\text{Rerata} : 0.648 + 0.005 + 0.650 + 0.868 + 0.738 + 0.005 + 0.547 + 0.321 + 0.811 + 0.648$$

$$10 : 0.5241$$

Rerata confidence Lama_studi dengan nilai TERLAMBAT

confidence(...
0.352
0.995
0.350
0.132
0.262
0.995
0.453
0.679
0.189
0.352

$$\text{Rerata} : 0.352 + 0.995 + 0.350 + 0.132 + 0.262 + 0.995 + 0.453 + 0.679 + 0.189 + 0.352$$

$$10 : 0.4759$$

Jumlah yang akan lulus TEPAT : 3 Orang

prediction(L...
TERLAMBAT
TEPAT
TERLAMBAT
TERLAMBAT
TERLAMBAT
TEPAT
TERLAMBAT
TEPAT
TERLAMBAT
TERLAMBAT

Jumlah yang akan lulus TERLAMBAT : 7 Orang

prediction(L)	
	TERLAMBAT
	TEPAT
	TERLAMBAT
	TERLAMBAT
	TERLAMBAT
	TEPAT
	TERLAMBAT
	TEPAT
	TERLAMBAT
	TERLAMBAT

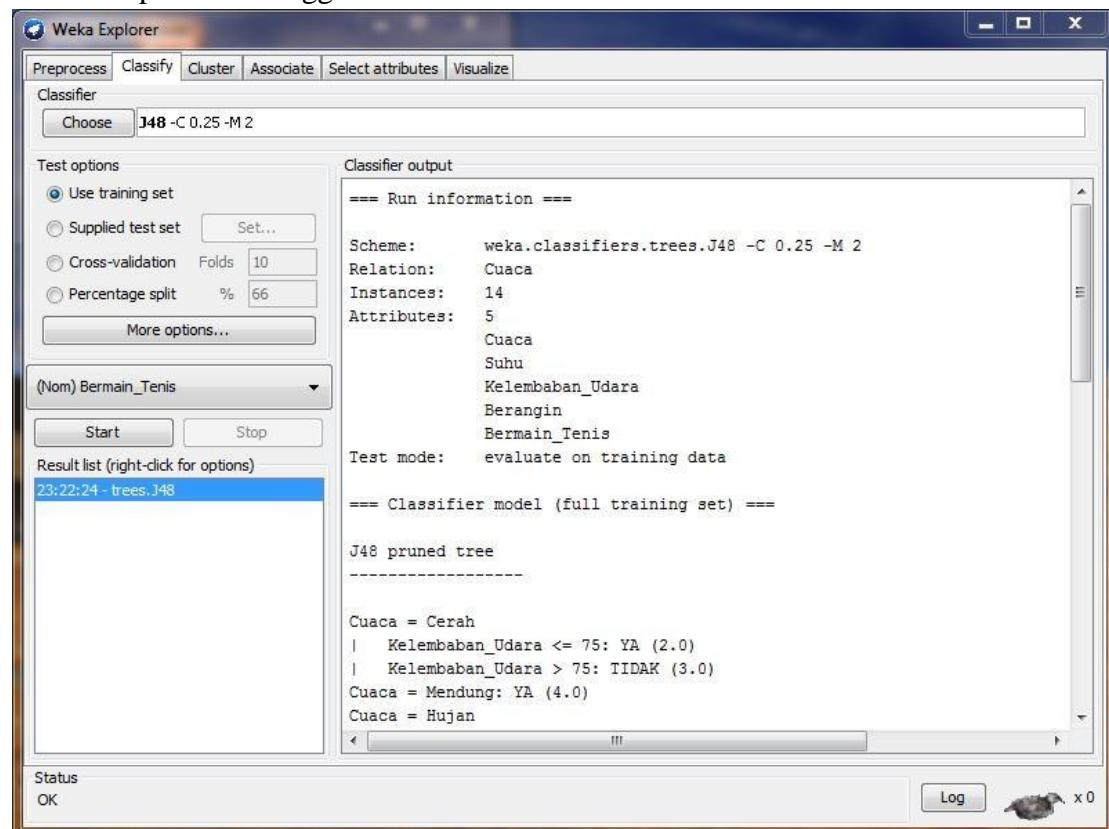
- ◆ Ketepatan lama studi Dewi, apabila Dewi adalah seorang WANITA dari jurusan IPA, asal sekolah LUAR SURAKARTA yang mengambil SKS dengan rata-rata 18 SKS tiap semester dan tidak pernah menjadi Asisten selama kuliah.
- ◆ Ketepatan lama studi Jono, apabila Jono adalah seorang PRIA dari jurusan selain IPA dan IPS, asal sekolah SURAKARTA yang mengambil SKS dengan rata-rata 17 SKS tiap semester dan pernah menjadi Asisten selama kuliah.
- ◆ Dewi dan Jono akan lulus TEPAT

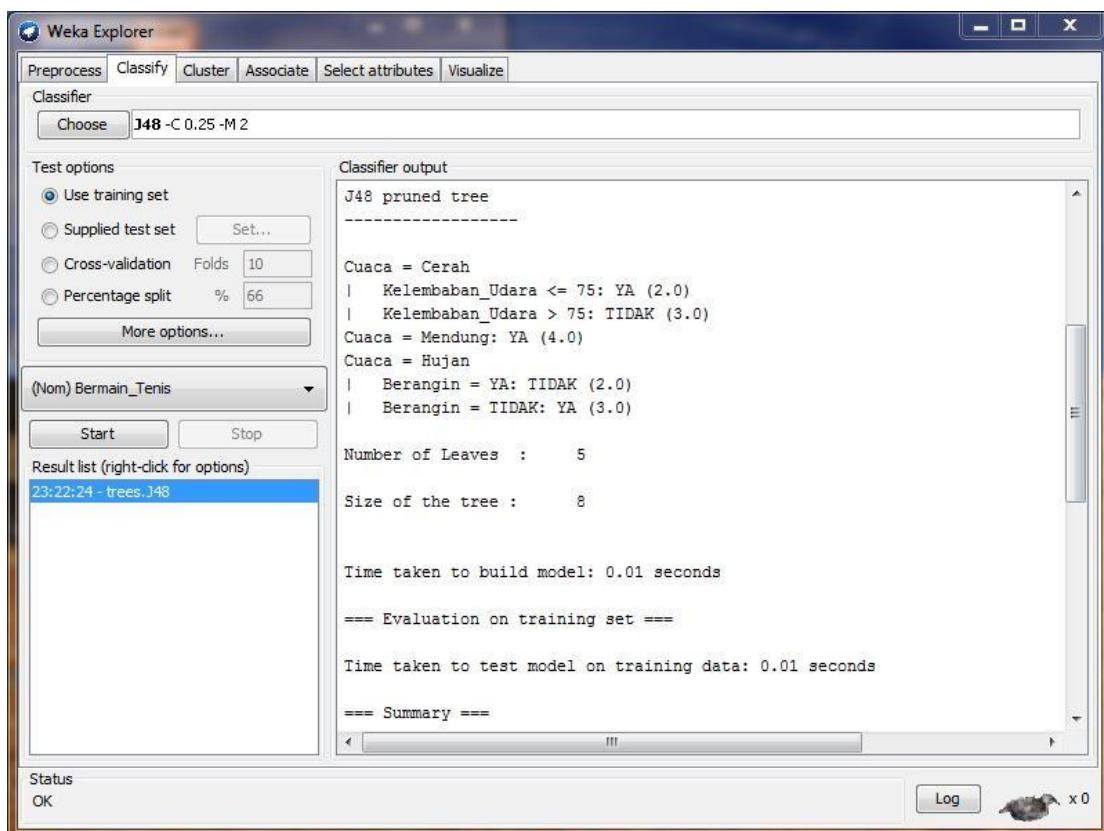
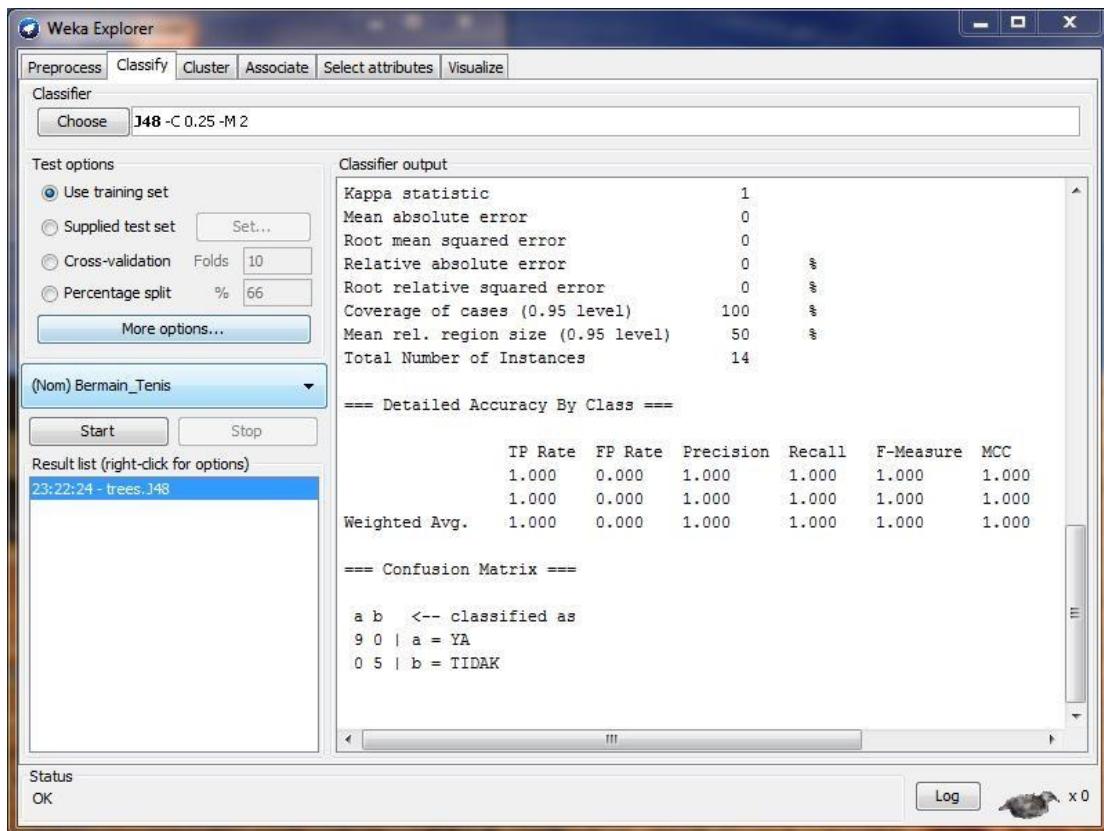
MODUL 9

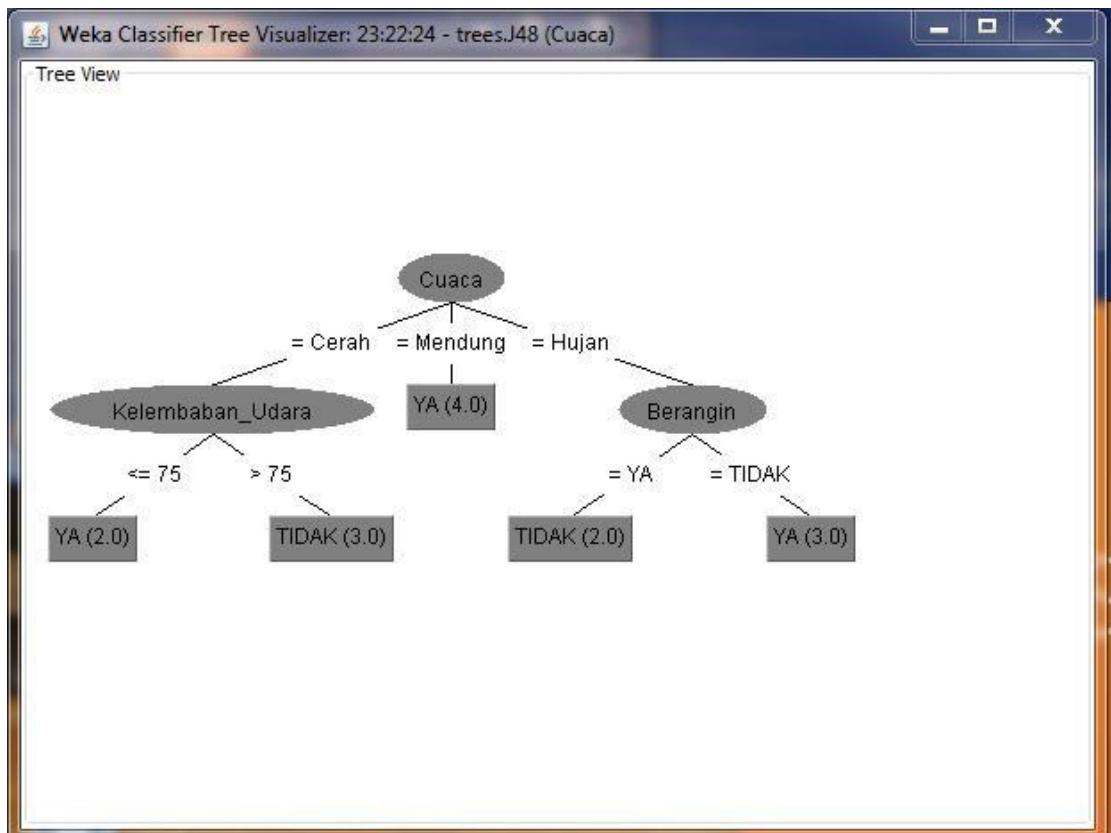
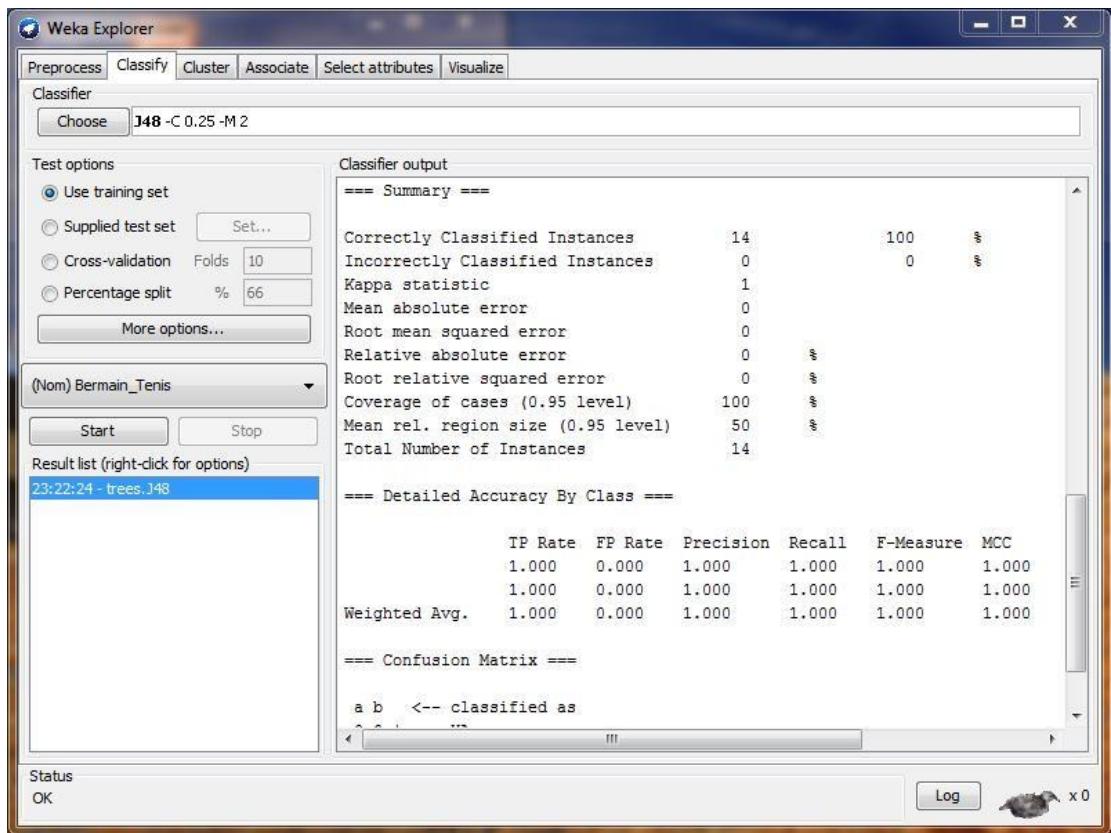
(Klasifikasi : Decision Tree)

KEGIATAN

Pohon Keputusan menggunakan Weka







Berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada sebagai berikut :

- | | |
|-----------------|----------------------------------|
| Simpul Akar | = Cuaca |
| Simpul internal | = Kelembapan_udara, dan Berangin |
| Simpul Daun | = YA, TIDAK |

Klasifikasi yang terbentuk yaitu :

Seseorang akan bermain tenis (YA) jika kondisi sebagai berikut :

Cuaca = Cerah, Kelembapan_udara ≤ 75 . (nilai atribut lain diabaikan)

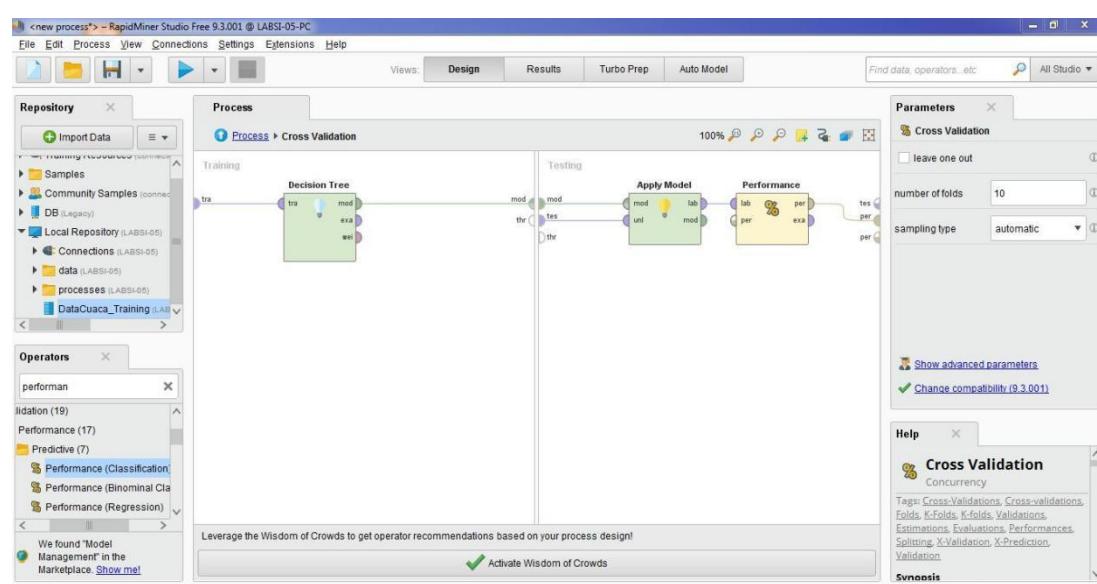
Cuaca = Mendung. (kondisi lain diabaikan)

Cuaca = Hujan, Berangin = TIDAK. (nilai atribut lain diabaikan)

Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :

Cuaca = Cerah, Kelembapan_udara > 75 . (nilai atribut lain diabaikan)

Cuaca = Hujan, Berangin = YA. (nilai atribut lain diabaikan)



Pohon Keputusan menggunakan Rapid Miner

RapidMiner Studio Free 9.3.001 @ LABSI-05-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Result History Example Set (//Local Repository/DataCuaca_Training)

Data Statistics Visualizations Annotations

Open In Turbo Prep Auto Model

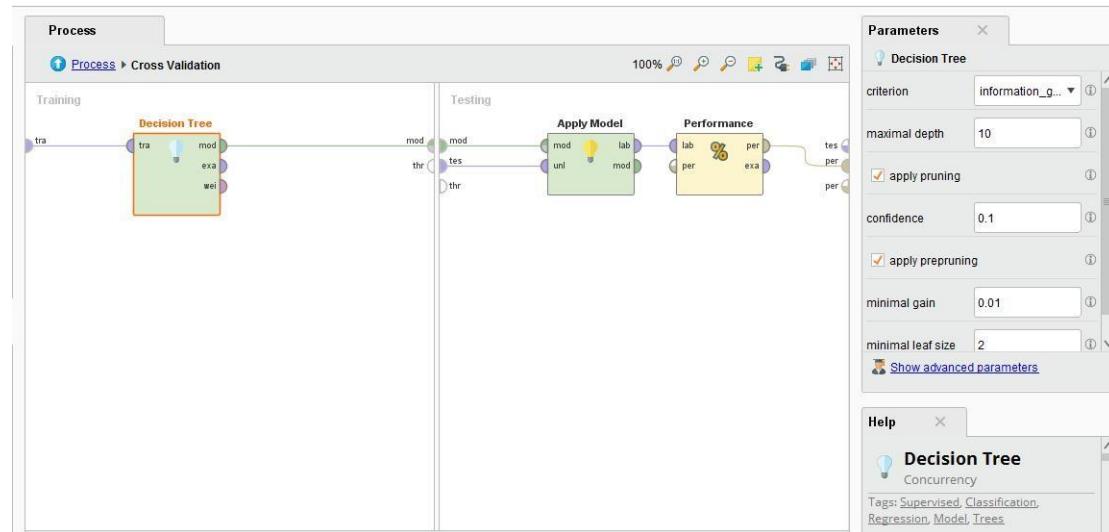
Filter (14 / 14 examples): all

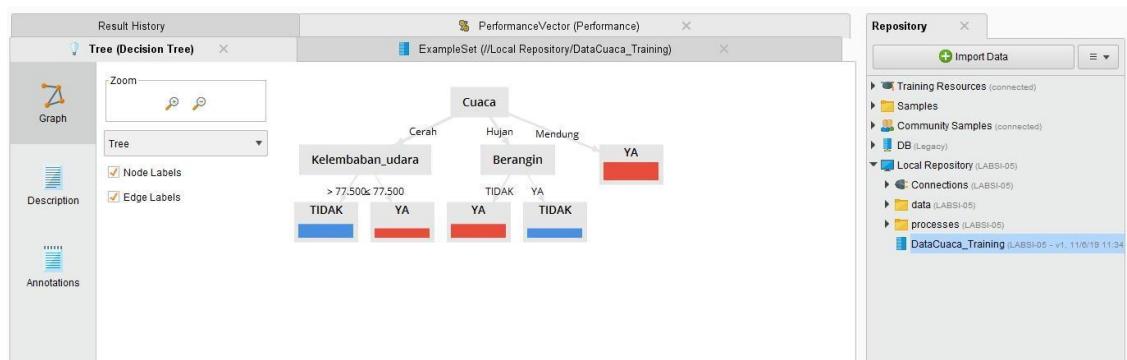
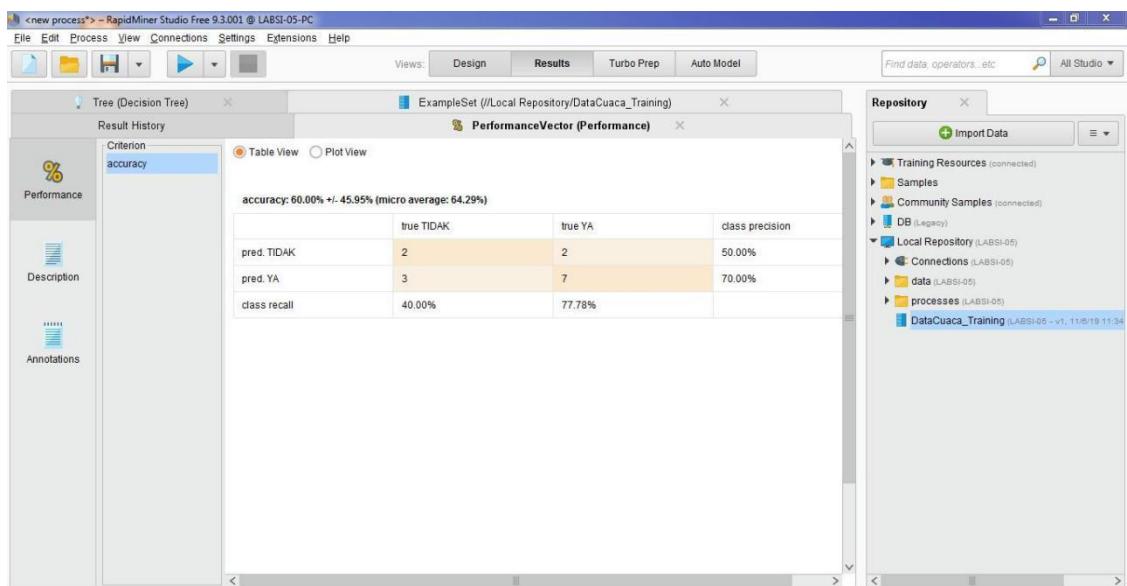
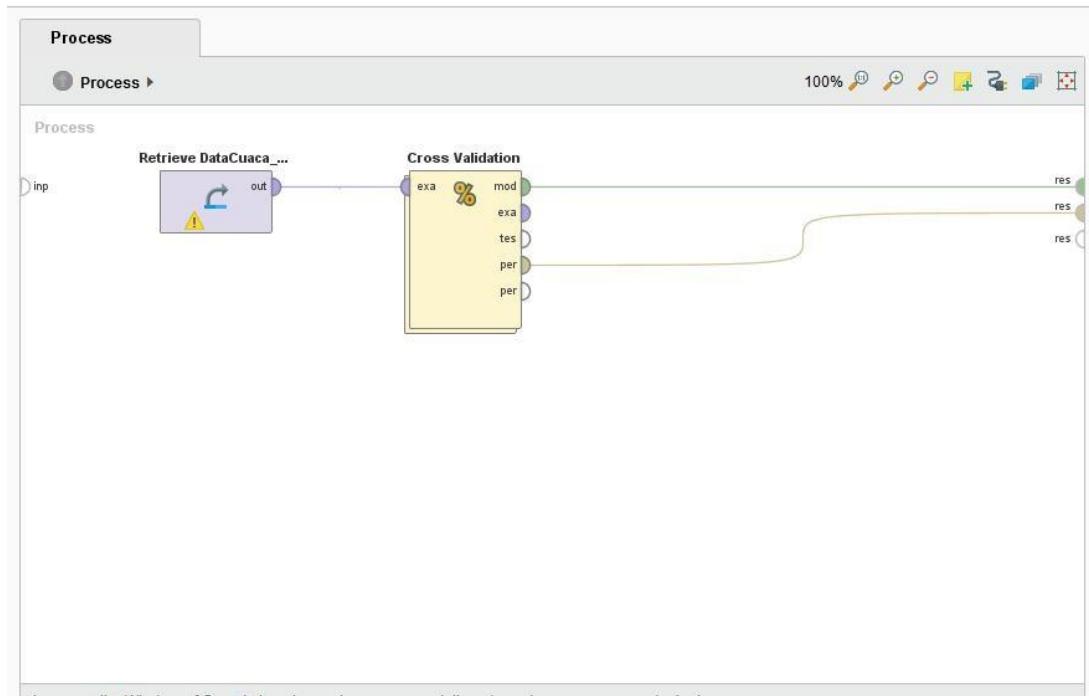
Row No.	Bermain_Te...	Cuaca	Suhu	Kelembaban...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	91	YA

ExampleSet (14 examples, 1 special attribute, 4 regular attributes)

Repository

- Import Data
- Training Resources (connected)
 - Samples
 - Community Samples (connected)
 - DB (Legacy)
 - Local Repository (LABSI-05)
 - Connections (LABSI-05)
 - data (LABSI-05)
 - processes (LABSI-05)
 - DataCuaca_Training (LABSI-05 - vt, 11/01/19 11:34)





Berdasarkan pohon keputusan tersebut, dapat dilihat jenis-jenis simpul yang ada sebagai berikut :

- | | |
|-----------------|----------------------------------|
| Simpul Akar | = Cuaca |
| Simpul internal | = Kelembapan_udara, dan Berangin |
| Simpul Daun | = YA, TIDAK |

Klasifikasi yang terbentuk yaitu :

Seseorang akan bermain tenis (YA) jika kondisi sebagai berikut :

Cuaca = Cerah, Kelembapan_udara <= 77,5. (nilai atribut lain diabaikan)

Cuaca = Mendung. (kondisi lain diabaikan)

Cuaca = Hujan, Berangin = TIDAK. (nilai atribut lain diabaikan)

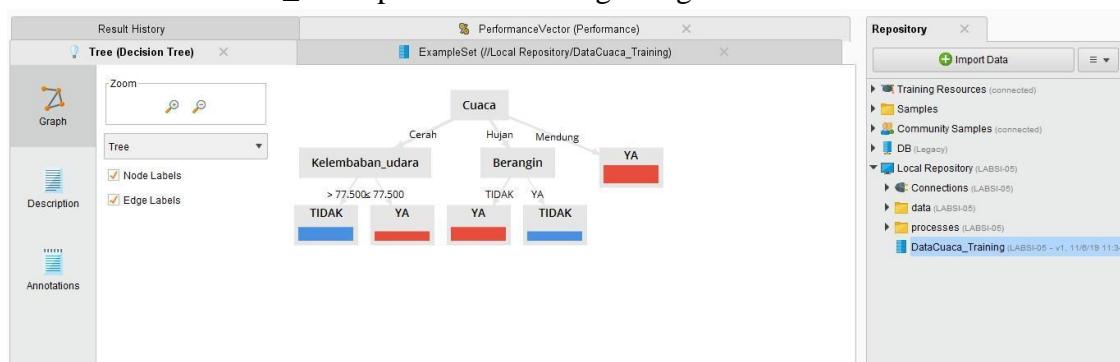
Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :

Cuaca = Cerah, Kelembapan_udara > 77,5. (nilai atribut lain diabaikan)

Cuaca = Hujan, Berangin = YA. (nilai atribut lain diabaikan)

Modul 9 (Klasifikasi : Decision Tree) TUGAS

Berdasarkan pohon keputusan pada kegiatan 9.4.2 (menggunakan Rapid Miner), nilai kelas atribut Bermain_Tenis pada tabel Testing sebagai berikut :



Cuaca	Suhu	Kelembapan_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

Berdasarkan pohon keputusan dari data Training pada Modul 7

Jumlah simpul daun pada pohon keputusan = 4

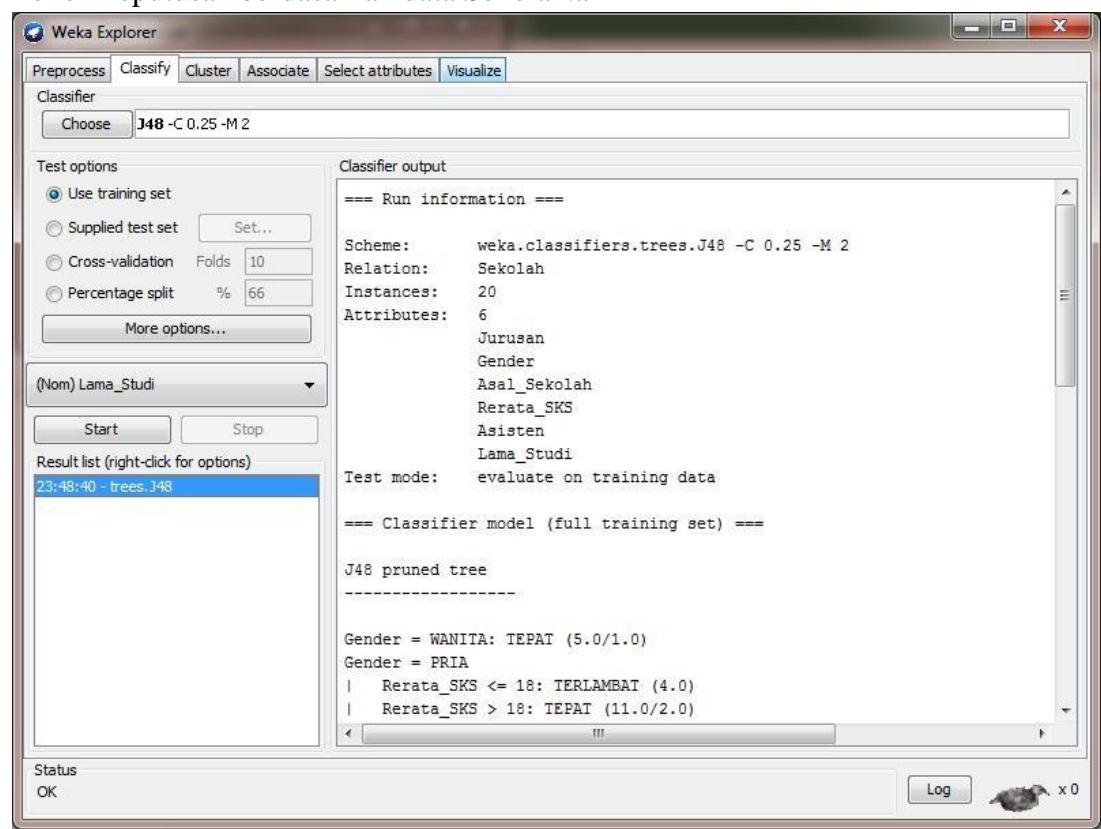
Jumlah simpul keseluruhan pada pohon keputusan = 7

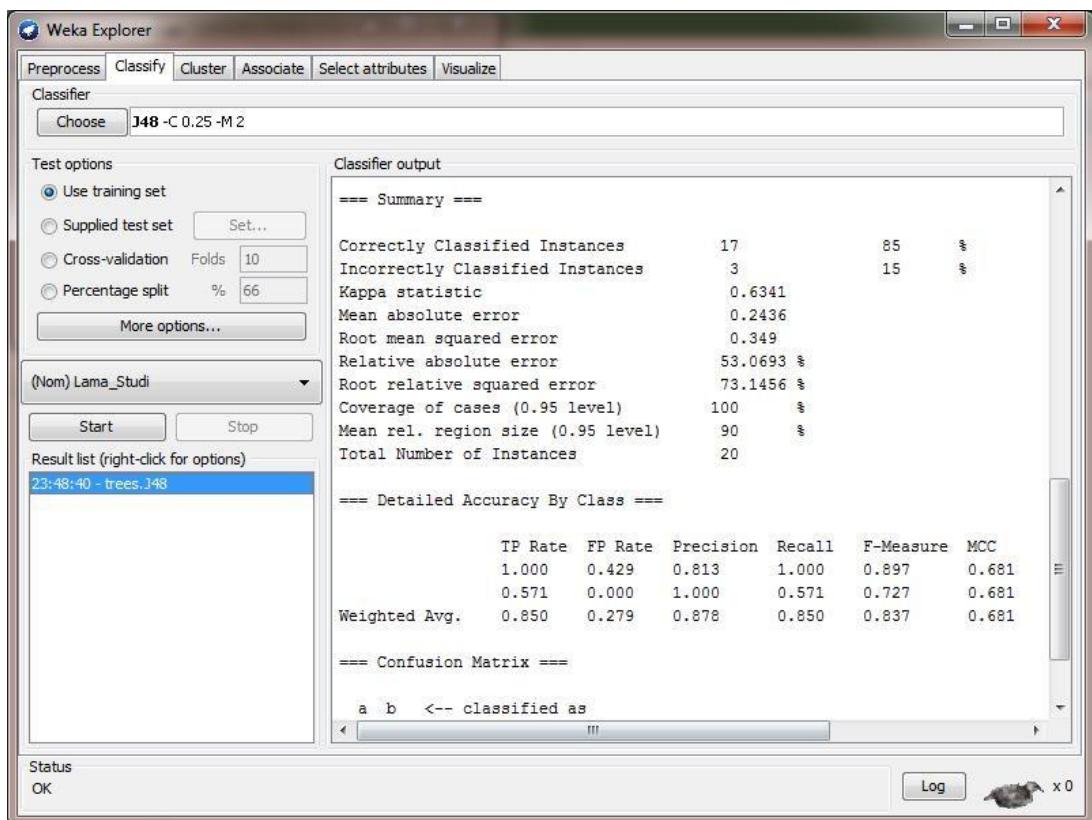
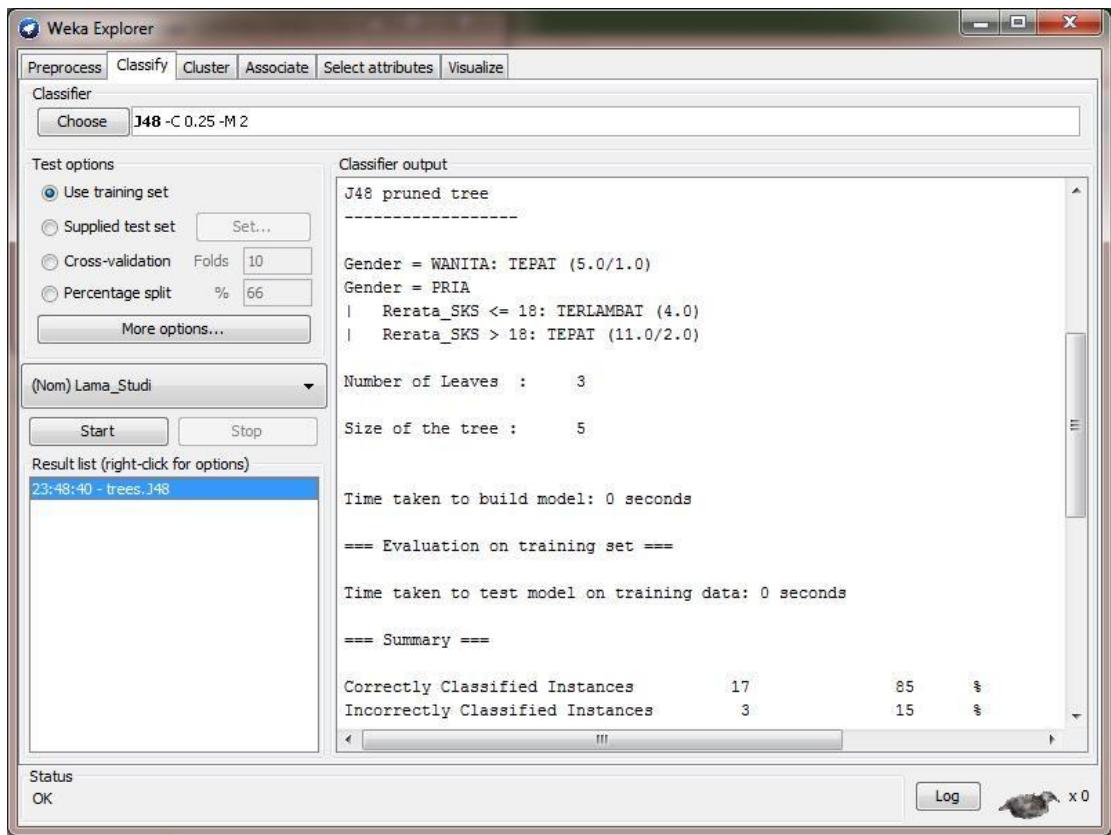
Waktu yang dibutuhkan untuk proses pelatihan = 0 detik

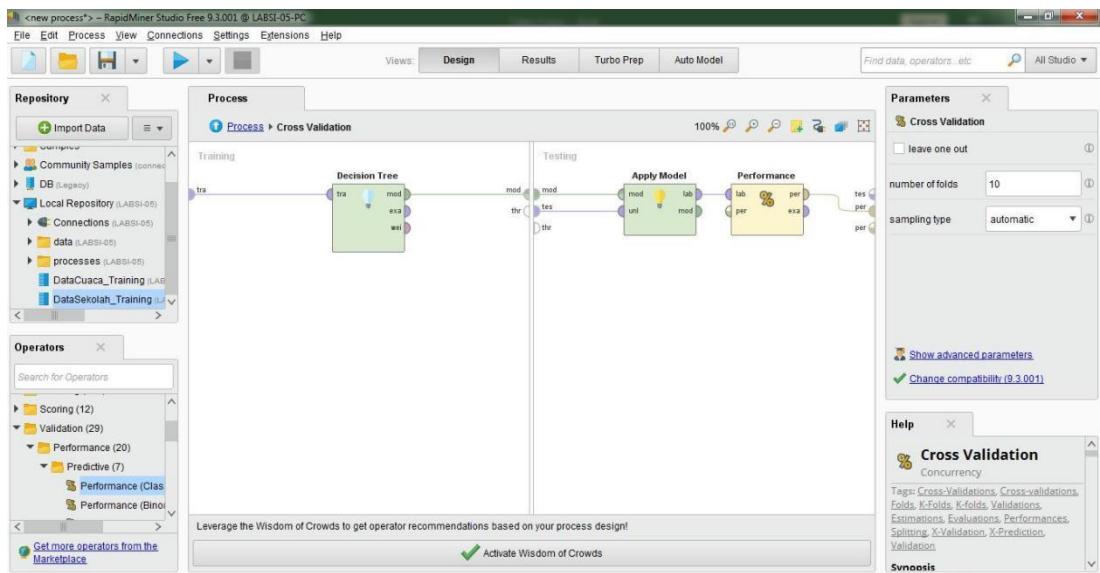
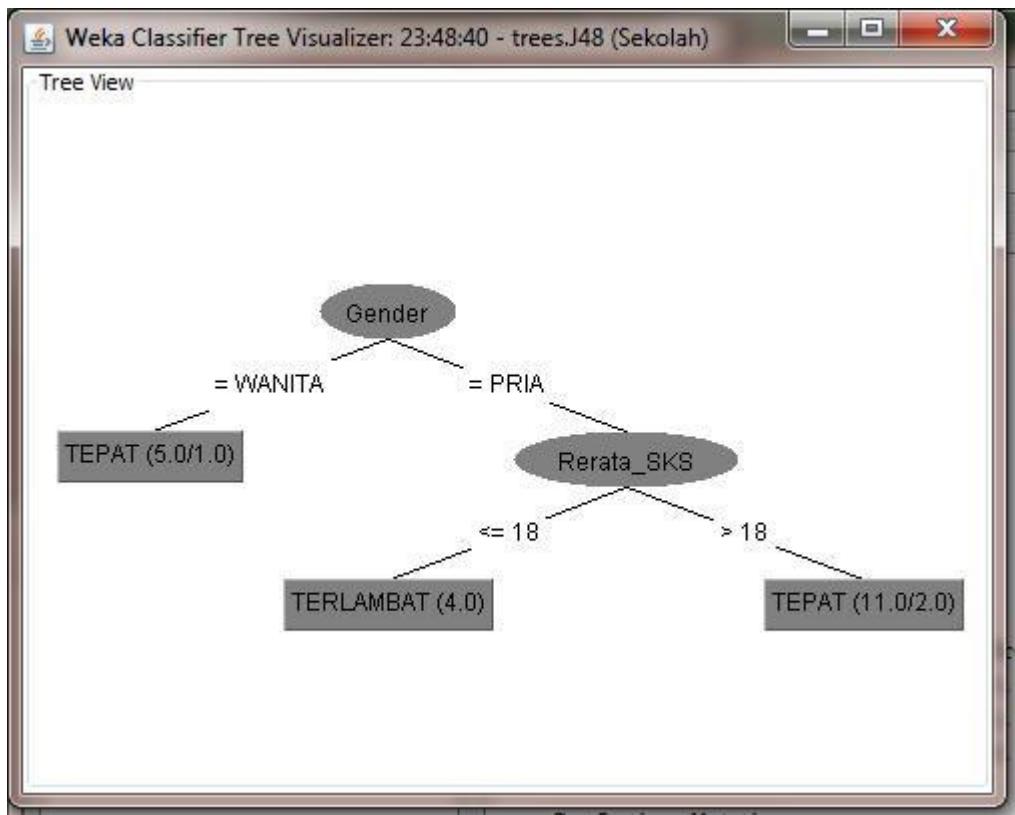
Tingkat ketepatan klasifikasi = 85%

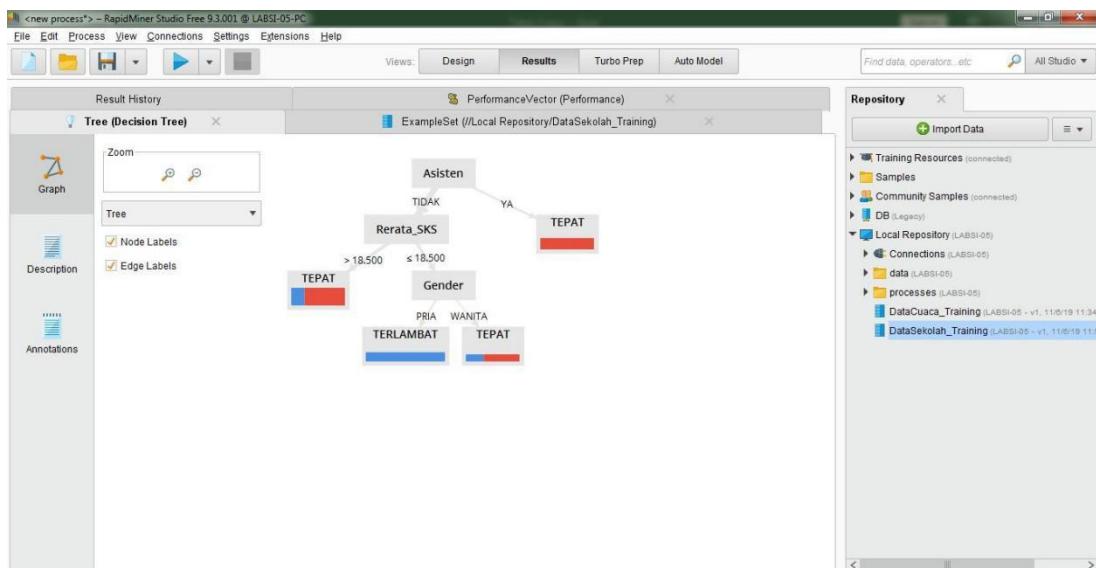
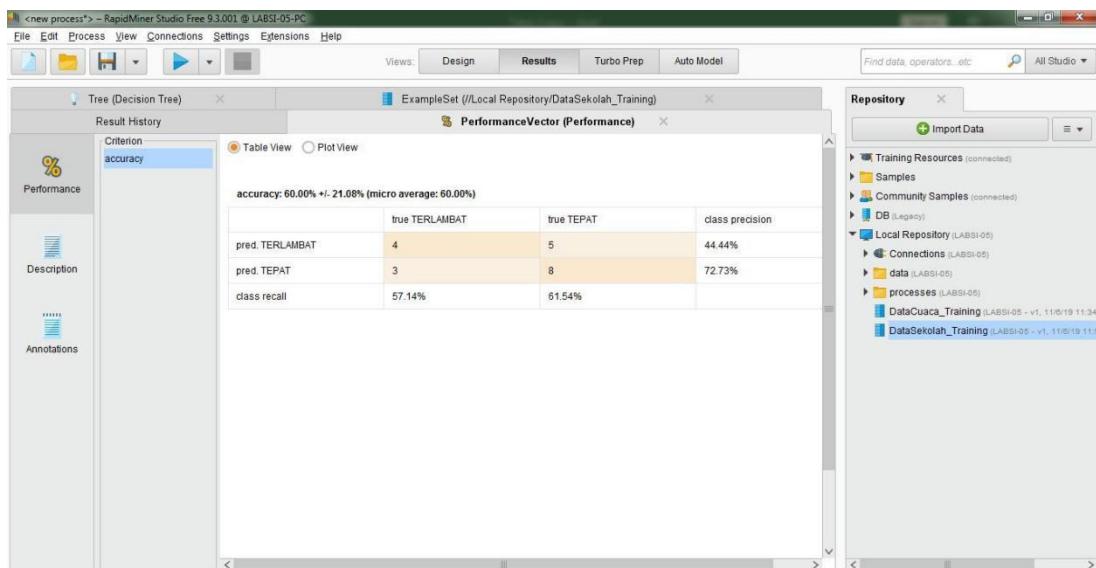
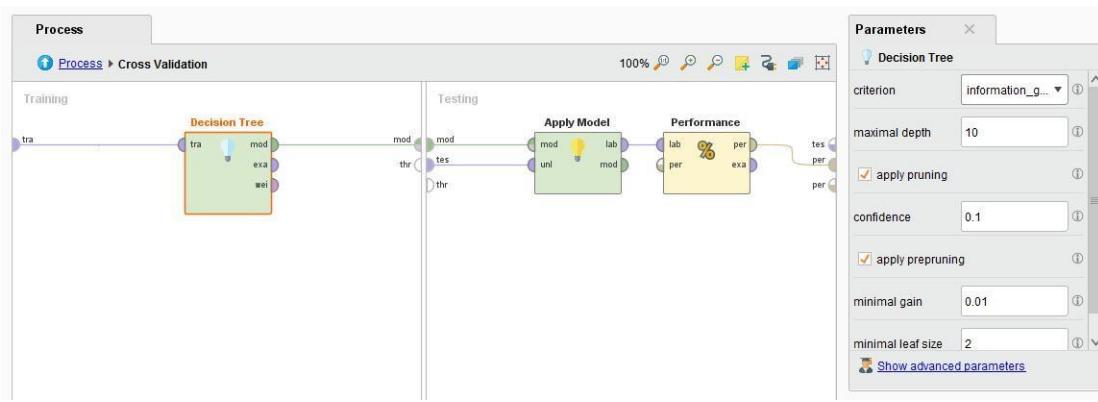
Tingkat ketidakakuratan klasifikasi = 15%

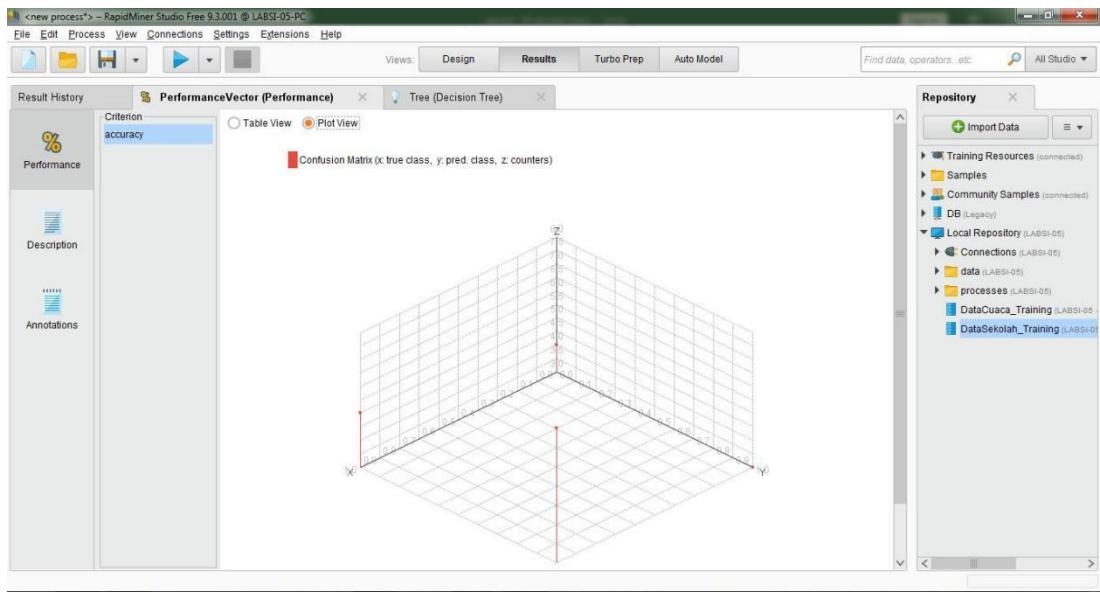
Pohon keputusan berdasarkan data Sekolah.arff











Jumlah simpul daun pada pohon keputusan = 3

Jumlah simpul keseluruhan pada pohon keputusan = 5

Waktu yang dibutuhkan untuk proses pelatihan = 0 detik

Tingkat ketepatan klasifikasi = 85%

Tingkat ketidaktepatan klasifikasi = 15%

MODUL 10

(Clustering : K-Means)

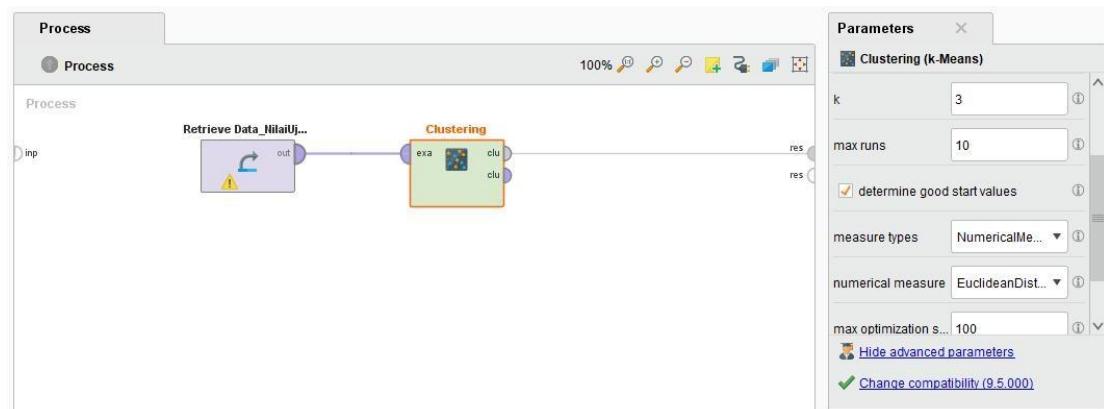
KEGIATAN

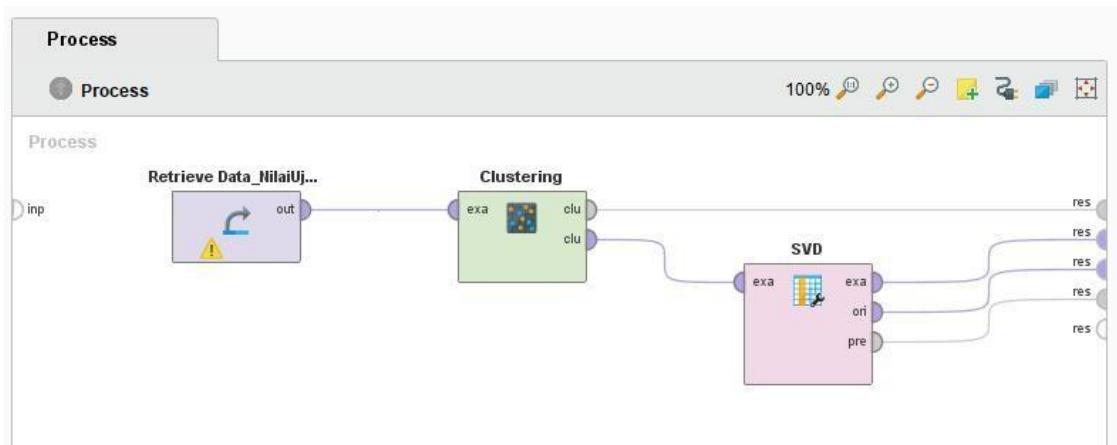
Tabel data nilai siswa

NO_SISWA	NAMA	B.IND	B.ING
S-101	JOKO	8.54	8.40
S-102	AGUS	9.98	6.81
S-103	SUSI	6.20	9.15
S-104	DYAH	5.24	7.26
S-105	WATI	5.70	5.71
S-106	IKA	8.57	5.87
S-107	EKO	7.70	7.71
S-108	YANTO	6.60	5.70
S-109	WAWAN	9.00	8.12
S-110	MAHMUD	9.81	9.58

	NAMA polynominal id	B.IND real	B.ING real
1	JOKO	8.540	8.400
2	AGUS	9.980	6.810
3	SUSI	6.200	9.150
4	DYAH	5.240	7.260
5	WATI	5.700	5.710
6	IKA	8.570	5.870
7	EKO	7.700	7.710
8	YANTO	6.600	5.700
9	WAWAN	9.000	8.120
10	MAHMUD	9.810	9.580

Area Design Process





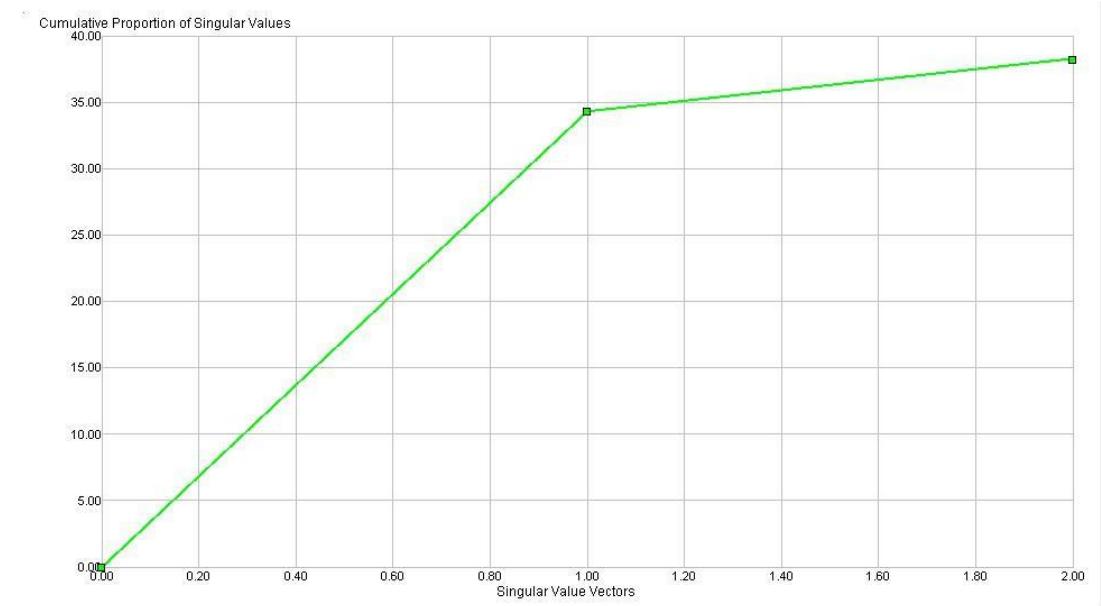
Hasil proses Clustering dengan Algoritma k-Means
 SVD (Singular Value Decomposition)
 Nilai Eigenvalue

Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Sin...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

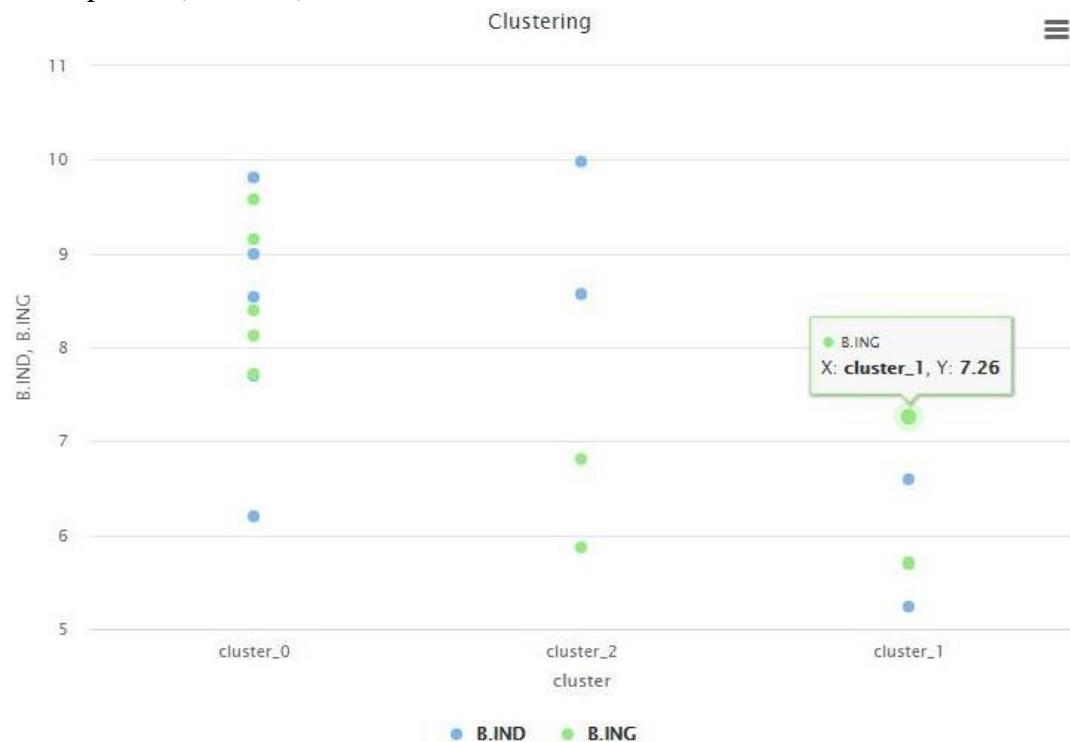
Nilai SVD Vectors

Attribute	SVD Vector 1
B.IND	0.723
B.ING	0.690

Nilai Cumulative Variance



ExampleSet (k-Means)



ExampleSet (SVD)

Row No.	NAMA	cluster ↑	svd_1
1	JOKO	cluster_0	0.349
3	SUSI	cluster_0	0.315
7	EKO	cluster_0	0.317
9	WAWAN	cluster_0	0.353
10	MAHMUD	cluster_0	0.399
4	DYAH	cluster_1	0.256
5	WATI	cluster_1	0.235
8	YANTO	cluster_1	0.254
2	AGUS	cluster_2	0.347
6	IKA	cluster_2	0.299

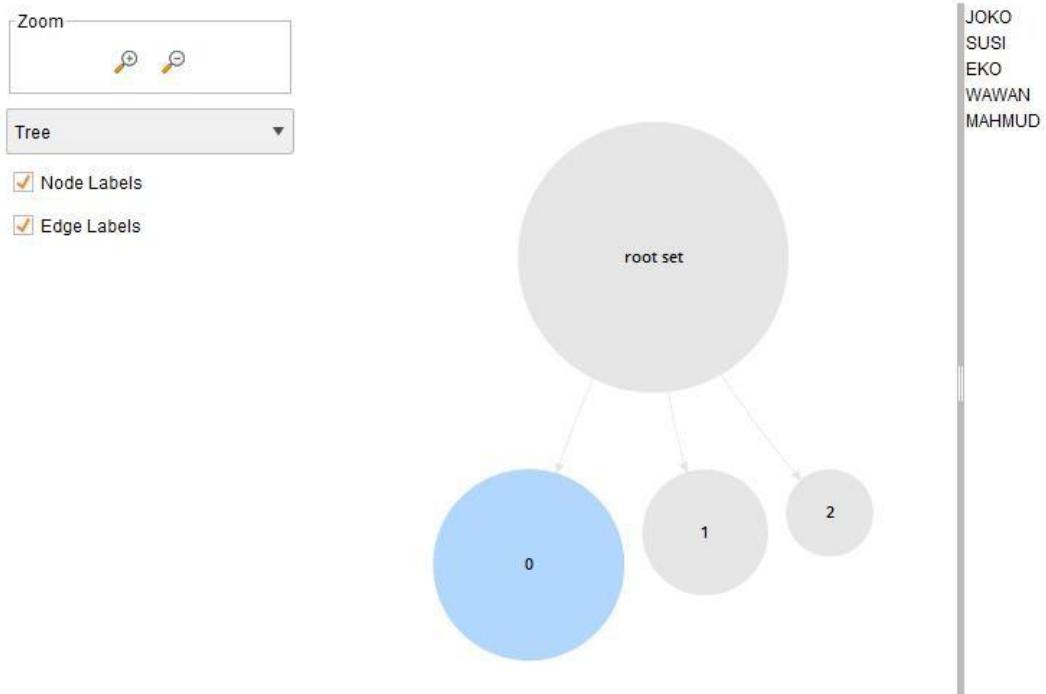
Cluster Model (Clustering)

Description

Cluster Model

```
Cluster 0: 5 items
Cluster 1: 3 items
Cluster 2: 2 items
Total number of items: 10
```

Graph



MODUL 10

(Clustering : K-Means)

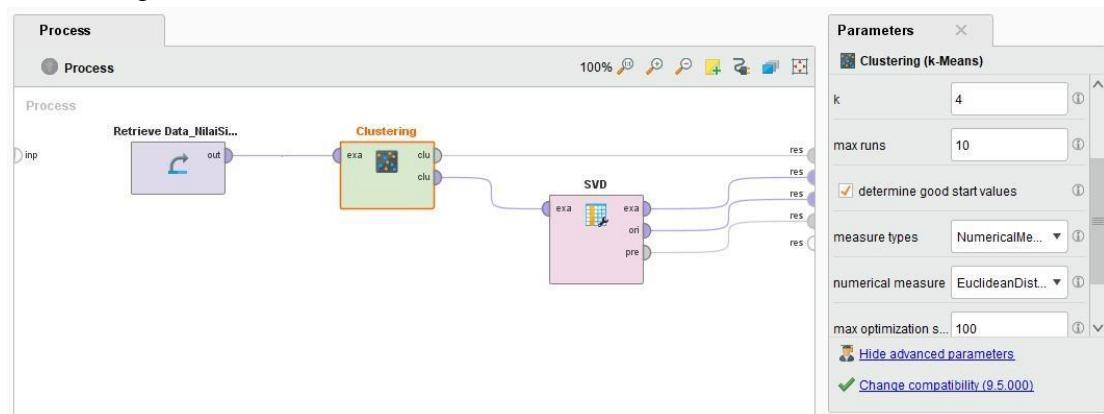
TUGAS

Tabel data nilai siswa

	NAMA polynominal id	B.IND real	B.ING real	MTK real	IPA real
1	JOKO	8.900	6.852	8.605	6.887
2	AGUS	5.498	8.557	8.480	6.810
3	SUSI	5.483	5.072	5.767	6.173
4	DYAH	8.663	6.160	9.301	5.023
5	WATI	6.937	9.437	5.735	9.007
6	IKA	8.390	7.450	8.234	8.494
7	EKO	6.082	5.341	7.973	7.357
8	YANTO	7.413	5.219	8.108	7.984
9	WAWAN	9.885	5.871	7.991	6.502
10	MAHMUD	5.580	5.222	6.844	9.192
11	BUDI	9.656	7.772	9.391	8.375
12	SANTI	9.382	8.575	7.977	7.154
13	DIAN	9.922	9.640	8.555	9.675

no problems.

Area Design Process



Hasil proses Clustering dengan Algoritma k-Means

SVD (Singular Value Decomposition)

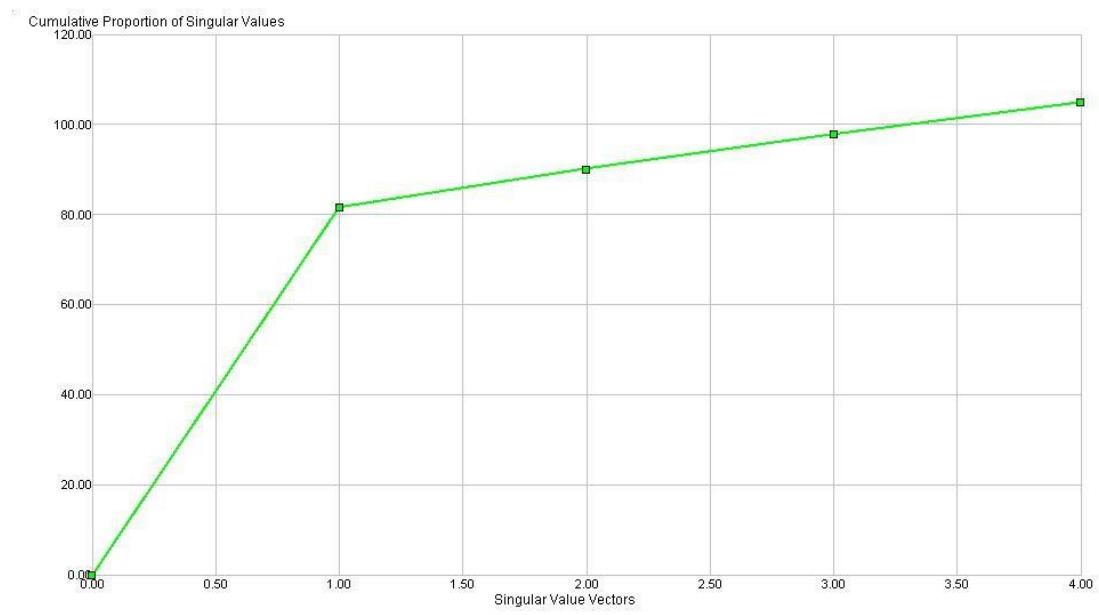
Nilai Eigenvalue

Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Sin...
SVD 1	81.661	0.778	81.661	0.778
SVD 2	8.538	0.081	90.199	0.859
SVD 3	7.708	0.073	97.906	0.933
SVD 4	7.081	0.067	104.987	1.000

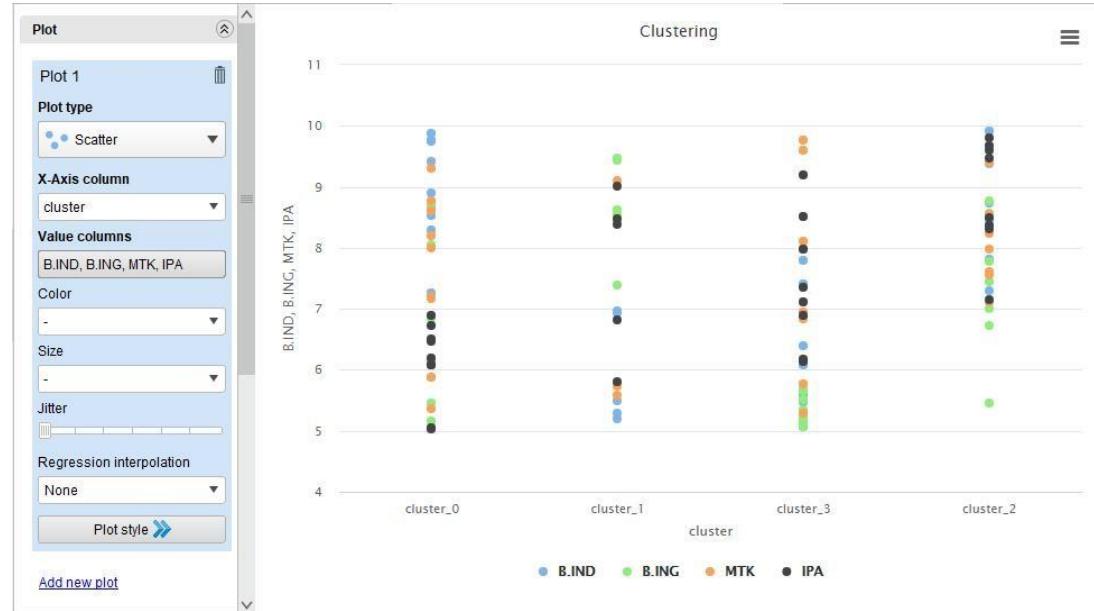
Nilai SVD Vectors

Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.IND	0.518	-0.552	0.587
B.ING	0.467	0.412	0.322
MTK	0.512	-0.406	-0.723
IPA	0.502	0.600	-0.170

Nilai Cumulative Variance



ExampleSet (k-Means)



ExampleSet (SVD)

Row No.	NAMA	cluster ↑	svd_1
1	JOKO	cluster_0	0.192
4	DYAH	cluster_0	0.179
9	WAWAN	cluster_0	0.186
15	AHMAD	cluster_0	0.178
17	RISA	cluster_0	0.176
19	YANI	cluster_0	0.162
22	JONO	cluster_0	0.193
23	SARAH	cluster_0	0.174
24	RAMA	cluster_0	0.176
2	AGUS	cluster_1	0.179
5	WATI	cluster_1	0.189
14	DANI	cluster_1	0.175
20	RATIH	cluster_1	0.174
30	TONI	cluster_1	0.175

Row No.	NAMA	cluster ↑	svd_1
6	IKA	cluster_2	0.200
11	BUDI	cluster_2	0.216
12	SANTI	cluster_2	0.202
13	DIAN	cluster_2	0.231
16	BAYU	cluster_2	0.190
21	INDAH	cluster_2	0.191
25	BAMBANG	cluster_2	0.194
28	FEBRI	cluster_2	0.211
3	SUSI	cluster_3	0.138
7	EKO	cluster_3	0.164
8	YANTO	cluster_3	0.177
10	MAHMUD	cluster_3	0.165
18	RANI	cluster_3	0.174
26	HADI	cluster_3	0.171
27	NANA	cluster_3	0.152
29	DENI	cluster_3	0.164

Cluster Model (Clustering)

Description

Cluster Model

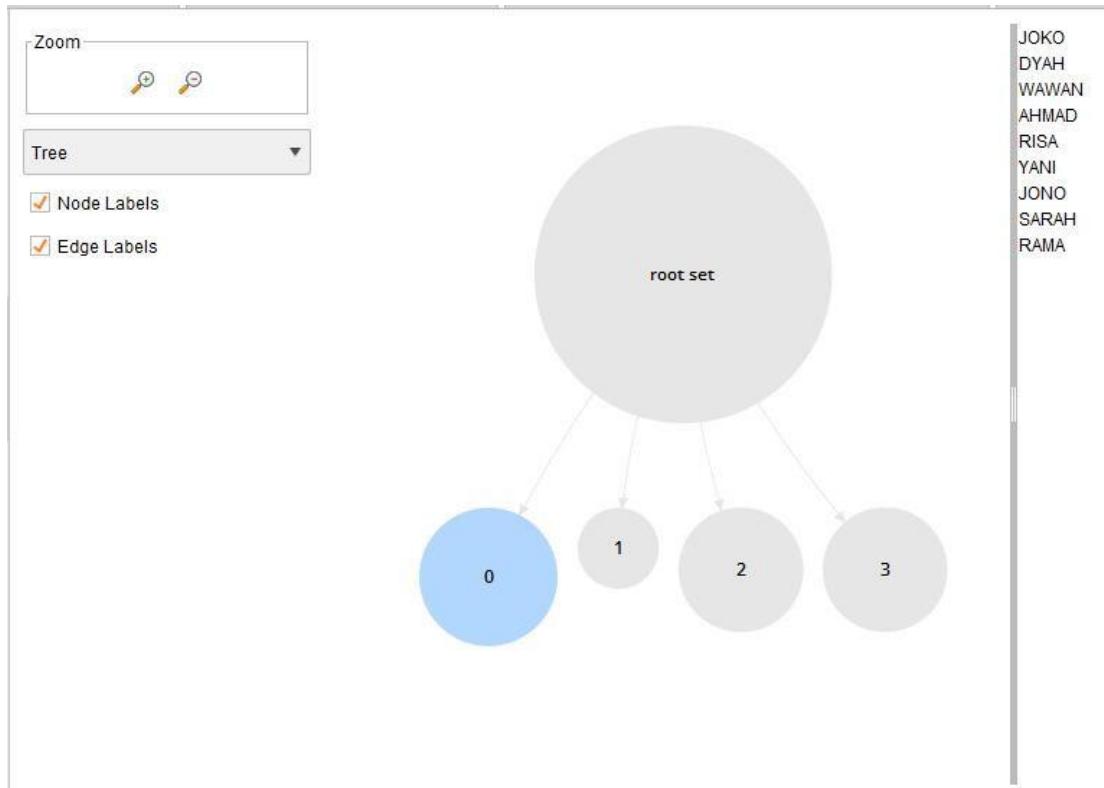
```

Cluster 0: 9 items
Cluster 1: 5 items
Cluster 2: 8 items
Cluster 3: 8 items
Total number of items: 30

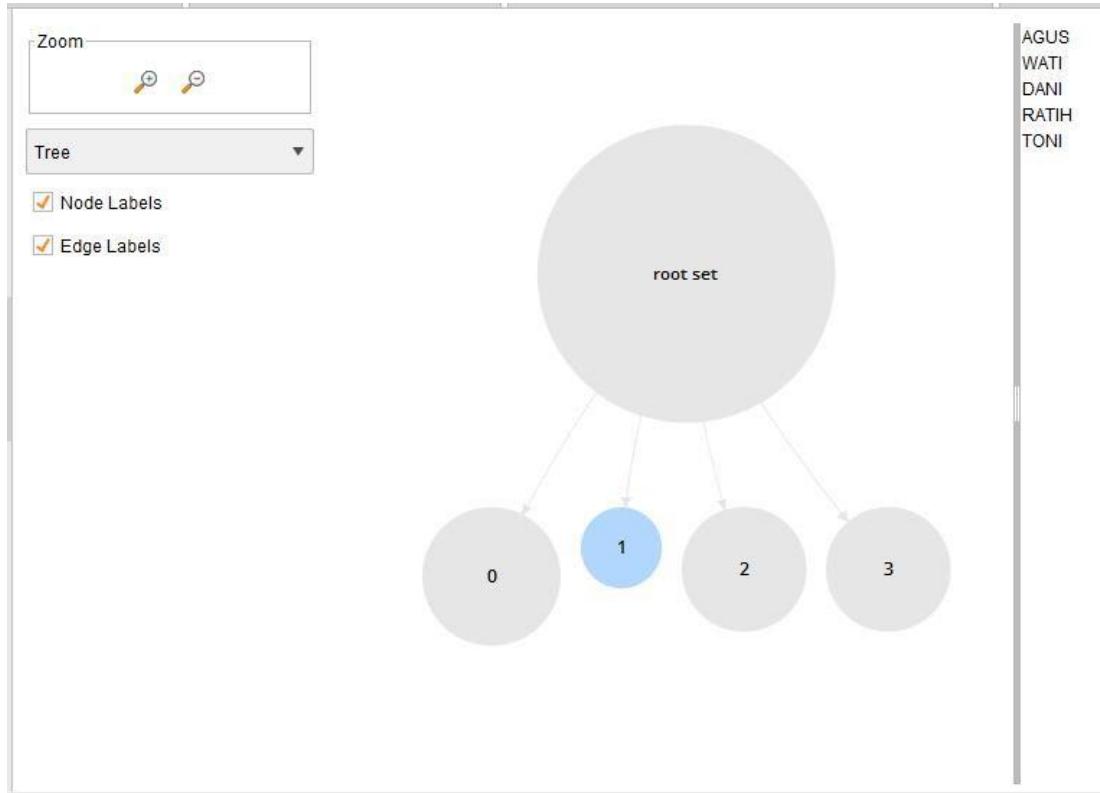
```

Graph

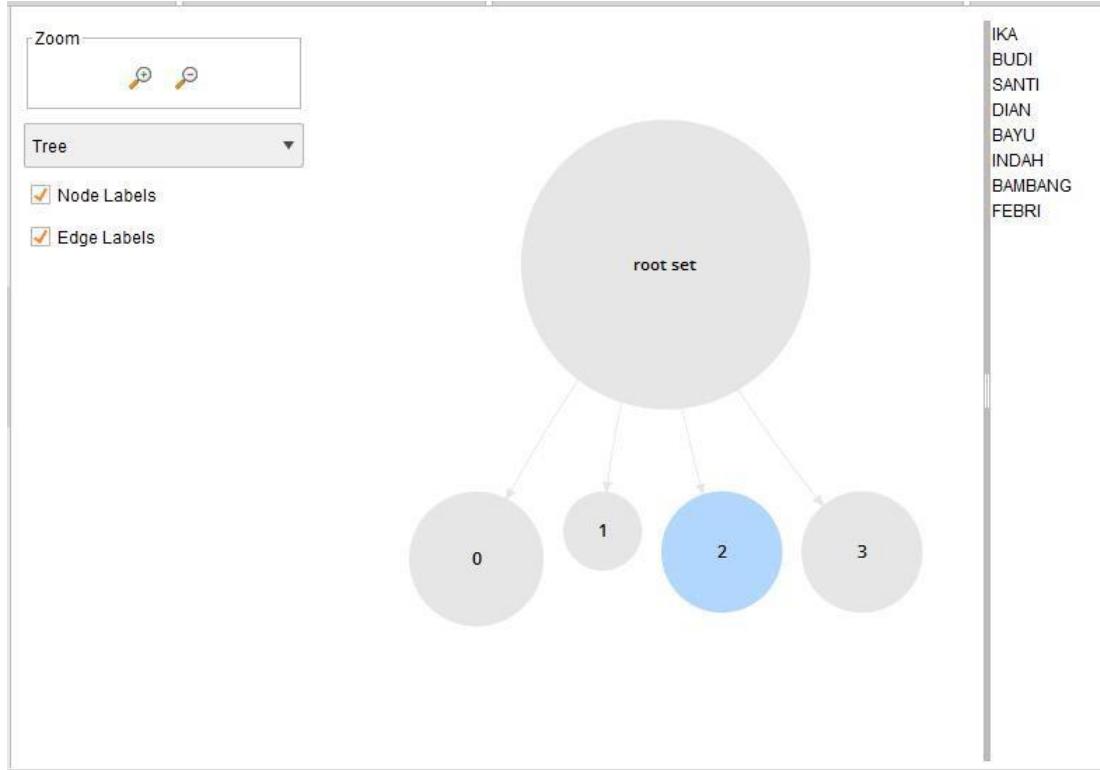
Cluster 0 : JOKO, DYAH, WAWAN, AHMAD, RISA, YANI, JONO, SARAH, RAMA



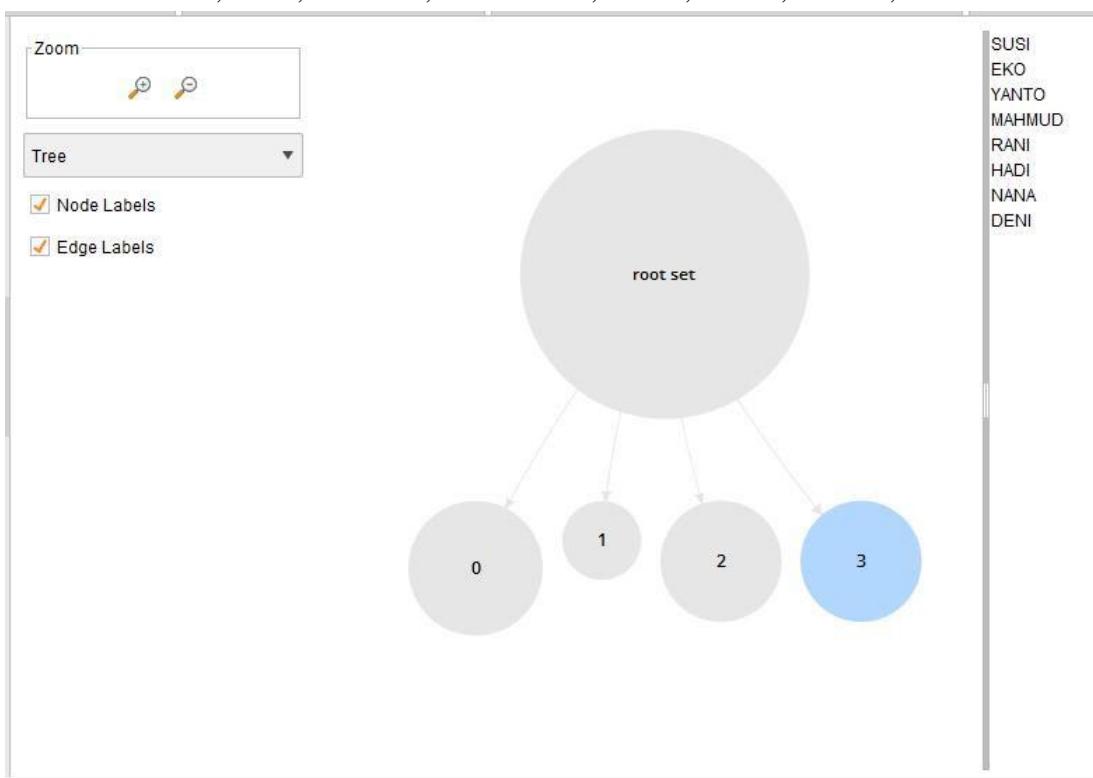
Cluster 1 : AGUS, WATI, DANI, RATIH, TONI



Cluster 2 : IKA, BUDI, SANTI, DIAN, BAYU, INDAH, BAMBANG, FEBRI

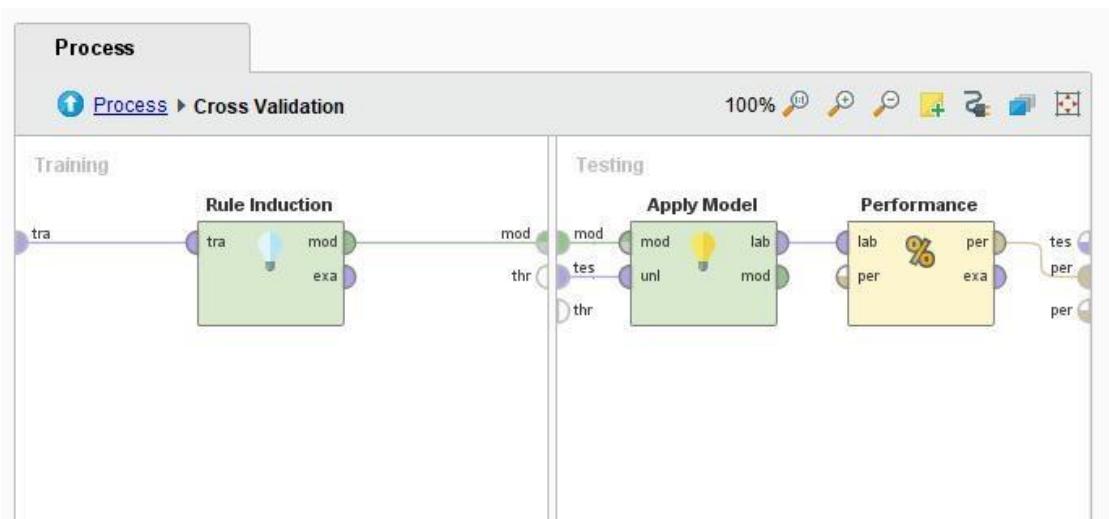
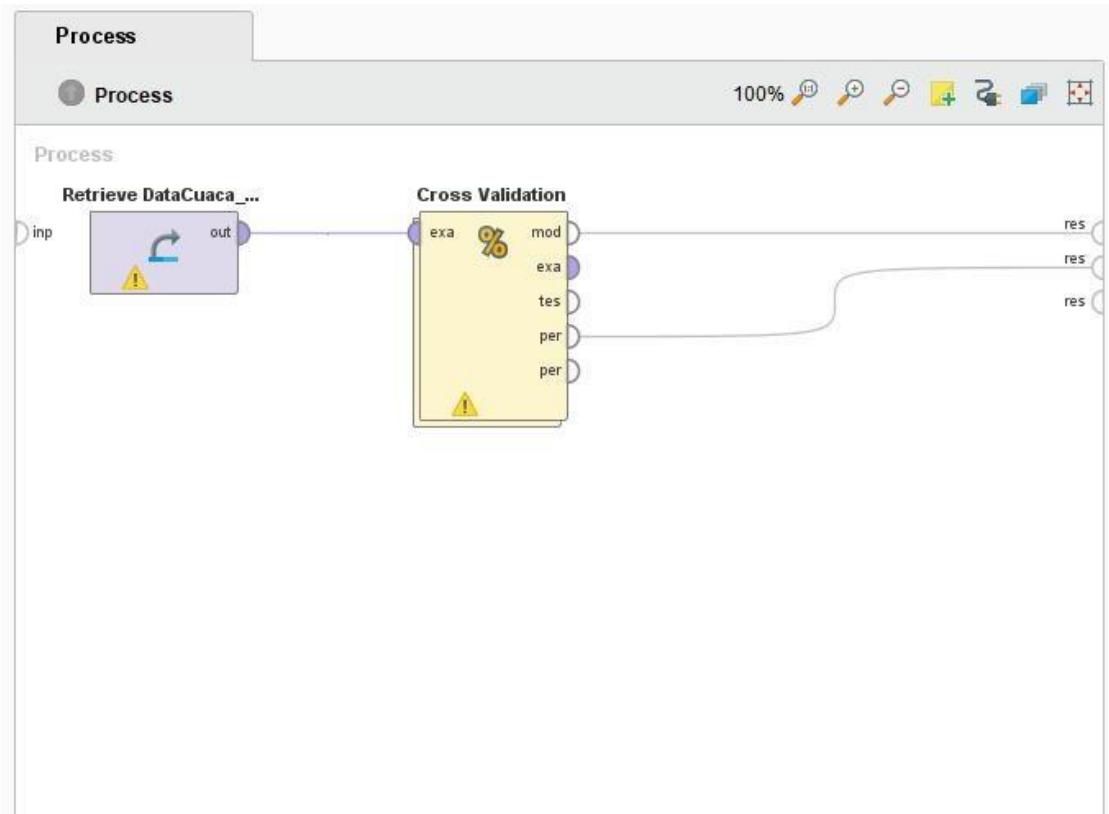


Cluster 3 : SUSI, EKO, YANTO, MAHMUD, RANI, HADI, NANA, DENI



MODUL 11
(Induksi dan Aturan Asosiasi)
KEGIATAN

Row No.	Bermain_Te...	Cuaca	Suhu	Kelembaban...	Berangin
1	TIDAK	Cerah	85	85	TIDAK
2	TIDAK	Cerah	80	90	YA
3	YA	Mendung	83	86	TIDAK
4	YA	Hujan	70	96	TIDAK
5	YA	Hujan	68	80	TIDAK
6	TIDAK	Hujan	65	70	YA
7	YA	Mendung	64	65	YA
8	TIDAK	Cerah	72	95	TIDAK
9	YA	Cerah	69	70	TIDAK
10	YA	Hujan	75	80	TIDAK
11	YA	Cerah	75	70	YA
12	YA	Mendung	72	90	YA
13	YA	Mendung	81	75	TIDAK
14	TIDAK	Hujan	71	91	YA



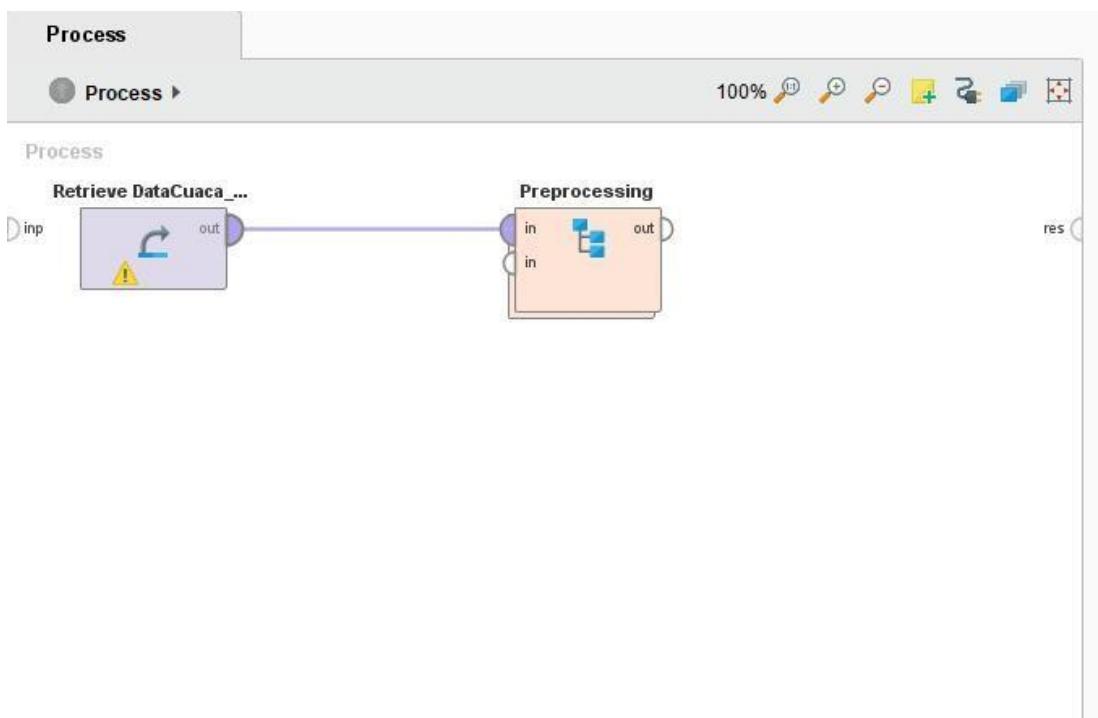
RuleModel

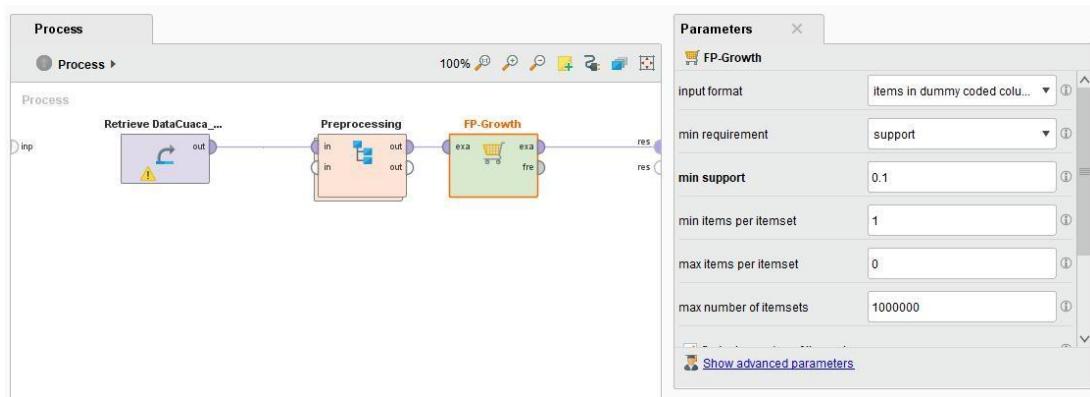
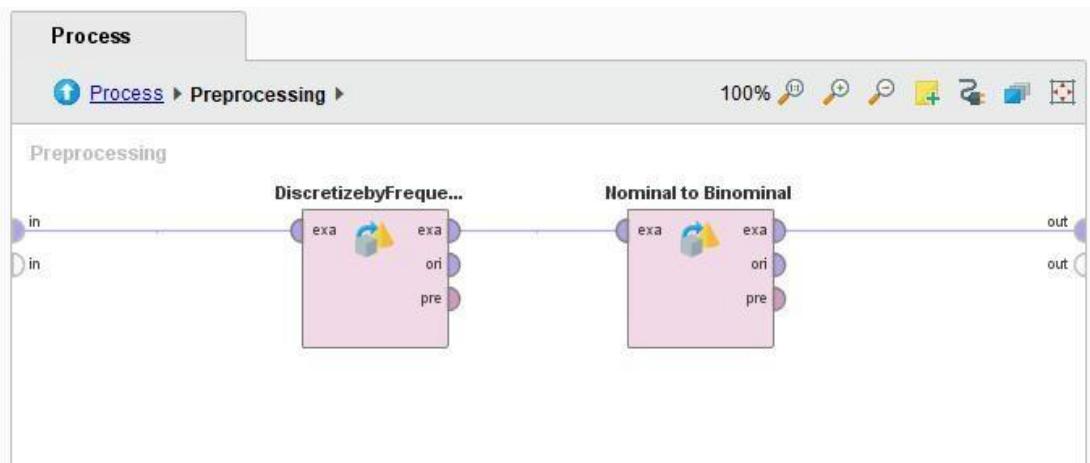
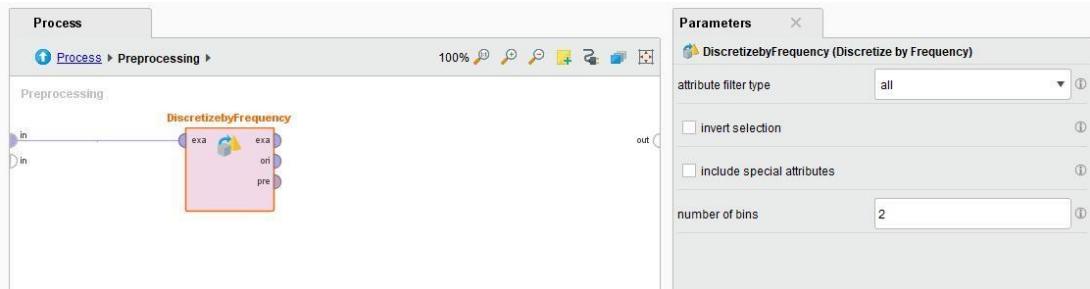
```
if Kelembaban_udara ≤ 82.500 then YA  (1 / 6)
if Cuaca = Cerah then TIDAK  (3 / 0)
if Cuaca = Mendung then YA  (0 / 2)
if Suhu ≤ 70.500 then YA  (0 / 1)
else TIDAK  (0 / 0)

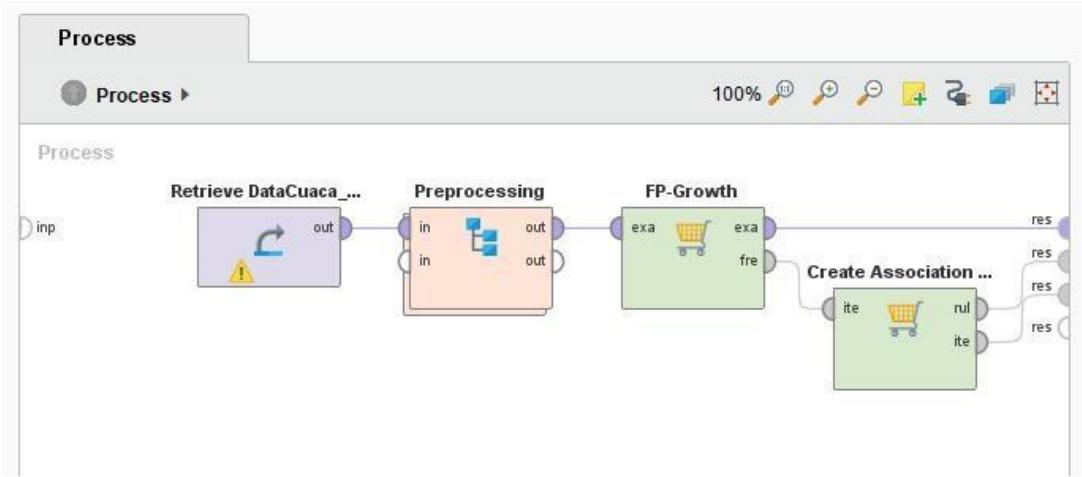
correct: 12 out of 13 training examples.
```

accuracy: 65.00% +/- 45.00% (micro average: 71.43%)

	true TIDAK	true YA	class precision
pred. TIDAK	2	1	66.67%
pred. YA	3	8	72.73%
class recall	40.00%	88.89%	







A. FREQUEN

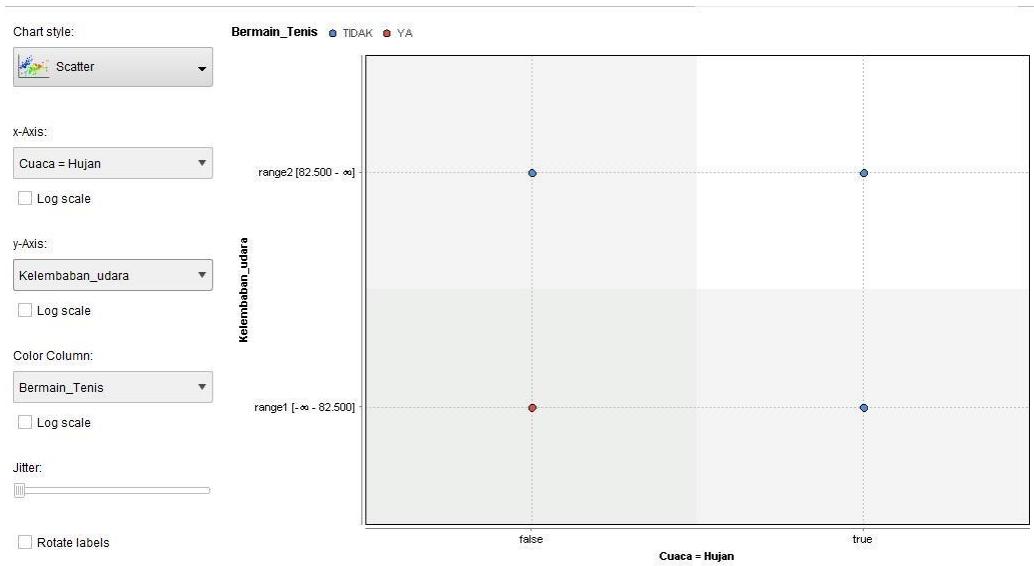
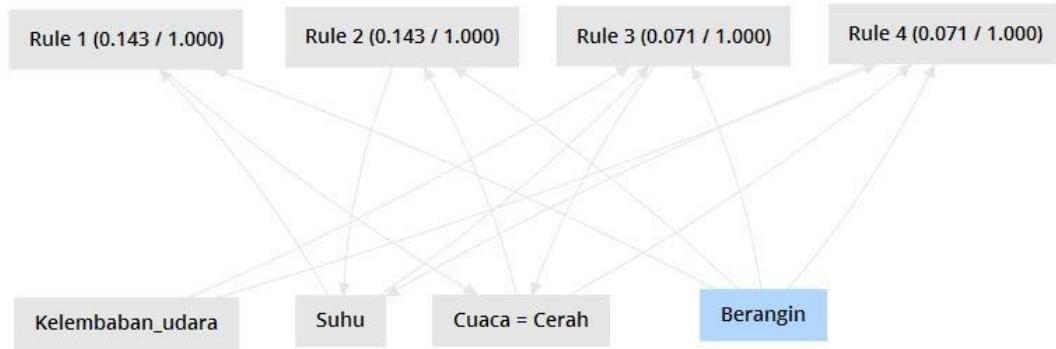
FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules) ExampleSet (Nominal to Binomial)

No. of Sets: 26 Total Max. Size: 4
Min. Size: 1 Max. Size: 4
Contains Item: Update View

Size	Support	Item 1	Item 2	Item 3	Item 4
1	0.500	Kelembaban_udara			
1	0.429	Berangin			
1	0.429	Suhu			
1	0.357	Cuaca = Cerah			
1	0.357	Cuaca = Hujan			
1	0.286	Cuaca = Mendung			
2	0.214	Kelembaban_udara	Berangin		
2	0.214	Kelembaban_udara	Suhu		
2	0.214	Kelembaban_udara	Cuaca = Cerah		
2	0.143	Kelembaban_udara	Cuaca = Hujan		
2	0.143	Kelembaban_udara	Cuaca = Mendung		
2	0.143	Berangin	Suhu		
2	0.143	Berangin	Cuaca = Cerah		
2	0.143	Berangin	Cuaca = Hujan		
2	0.143	Berangin	Cuaca = Mendung		
2	0.214	Suhu	Cuaca = Cerah		

No.	Premises	Conclusion	Support	Confidence
1	Berangin, Suhu	Cuaca = Cerah	0.143	1
2	Berangin, Cuaca = Cerah	Suhu	0.143	1
3	Kelembaban_udara, Berangin, Suhu	Cuaca = Cerah	0.071	1
4	Kelembaban_udara, Berangin, Cuaca = Cerah	Suhu	0.071	1

ession	Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
= Cerah	0.143	1	1	-0.143	0.092	2.800	∞
	0.143	1	1	-0.143	0.082	2.333	∞
= Cerah	0.071	1	1	-0.071	0.046	2.800	∞
	0.071	1	1	-0.071	0.041	2.333	∞



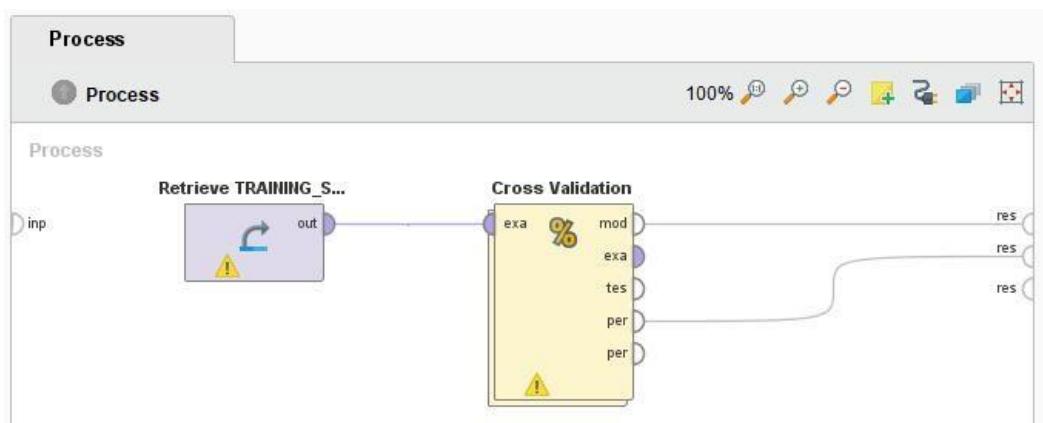
MODUL 11

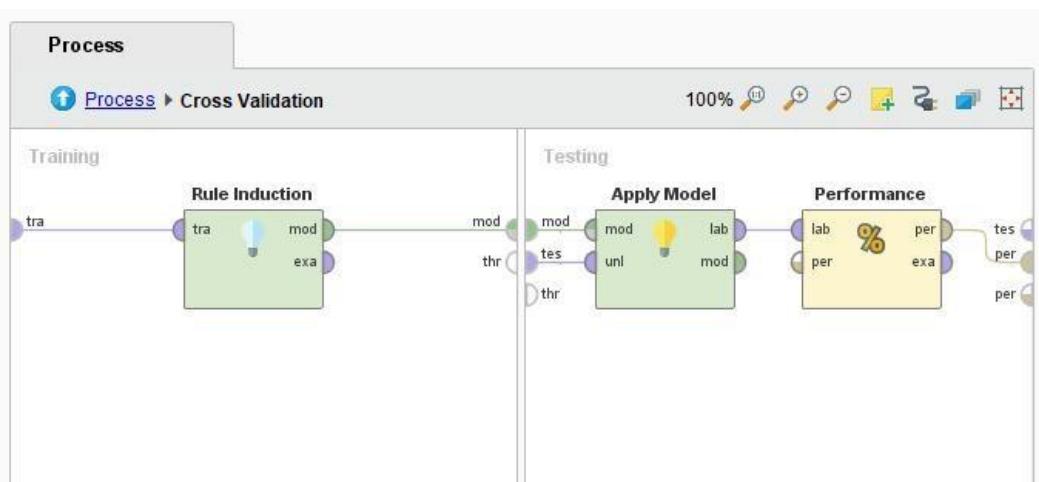
(Induksi dan Aturan Asosiasi)

TUGAS

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

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
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
15	TEPAT	IPA	PRIA	LUAR	22	YA
16	TERLAMBAT	LAIN	PRIA	SURAKARTA	16	TIDAK
17	TEPAT	IPS	PRIA	LUAR	20	TIDAK
18	TEPAT	LAIN	PRIA	LUAR	23	YA
19	TEPAT	IPA	PRIA	SURAKARTA	21	YA
20	TERLAMBAT	IPS	PRIA	SURAKARTA	19	TIDAK





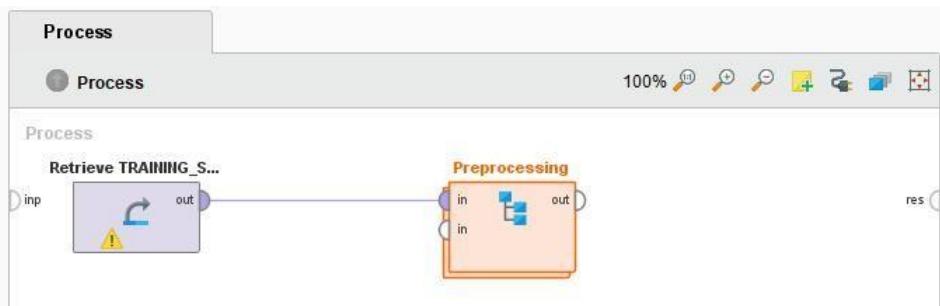
RuleModel

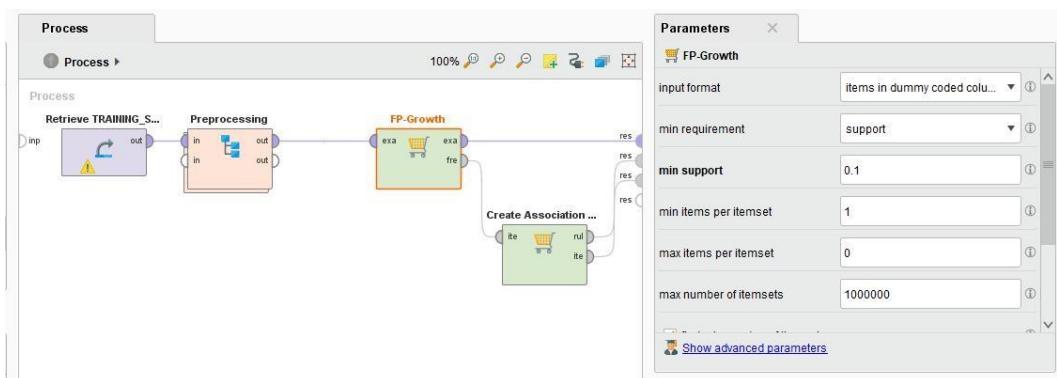
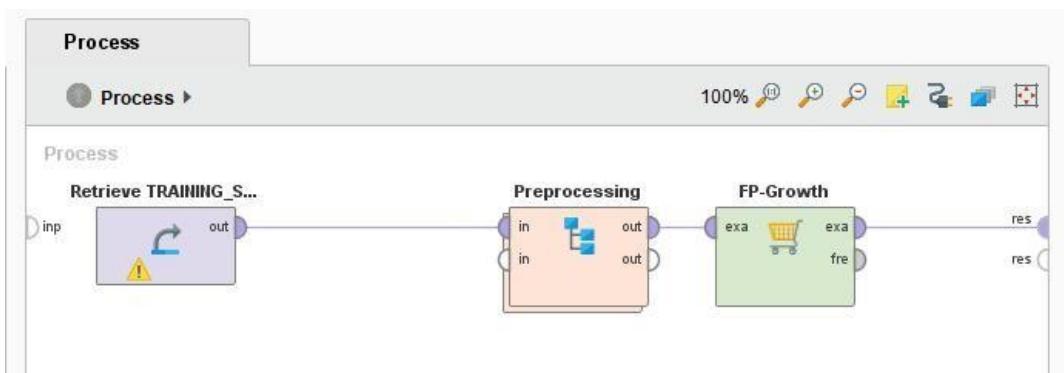
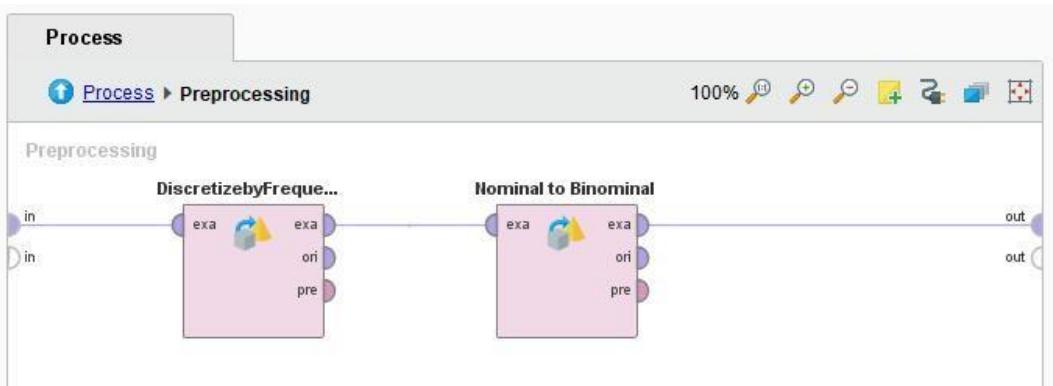
```
if Rerata_SKS > 18.500 then TEPAT (2 / 10)
if Gender = PRIA then TERLAMBAT (4 / 0)
if Jurusan_SMA = IPA then TEPAT (0 / 2)
if Jurusan_SMA = IPS then TERLAMBAT (1 / 0)
else TEPAT (0 / 0)
```

Table View Plot View

accuracy: 65.00% +/- 32.02% (micro average: 65.00%)

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	4	4	50.00%
pred. TEPAT	3	9	75.00%
class recall	57.14%	69.23%	





Result History FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules) ExampleSet (Nominal to Binomial)

Data

No. of Sets: 55	Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
Total Max. Size: 5	1	0.750	Gender				
Min. Size: 1	1	0.500	Jurusan_SMA = IPA				
Max. Size: 5	1	0.300	Asal_Sekolah				
Contains Item:	1	0.300	Jurusan_SMA = IPS				
	1	0.250	Asisten				
Update View	1	0.250	Rerata_SKS				
	1	0.200	Jurusan_SMA = L...				
	2	0.350	Gender	Jurusan_SMA = IPA			
	2	0.250	Gender	Asal_Sekolah			
	2	0.250	Gender	Jurusan_SMA = IPS			
	2	0.200	Gender	Asisten			
	2	0.250	Gender	Rerata_SKS			
	2	0.150	Gender	Jurusan_SMA = L...			
	2	0.150	Jurusan_SMA = IPA	Asal_Sekolah			
	2	0.200	Jurusan_SMA = IPA	Asisten			
	2	0.100	Jurusan_SMA = IPA	Rerata_SKS			

FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules) ExampleSet (Nominal to Binomial)

Data

No. of Sets: 55	Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
Total Max. Size: 5	2	0.100	Asal_Sekolah	Jurusan_SMA = IPS			
Min. Size: 1	2	0.150	Asal_Sekolah	Asisten			
Max. Size: 5	2	0.150	Asal_Sekolah	Rerata_SKS			
Contains Item:	2	0.050	Asal_Sekolah	Jurusan_SMA = L...			
	2	0.100	Jurusan_SMA = IPS	Rerata_SKS			
Update View	2	0.150	Asisten	Rerata_SKS			
	2	0.050	Asisten	Jurusan_SMA = L...			
	2	0.050	Rerata_SKS	Jurusan_SMA = L...			
	3	0.100	Gender	Jurusan_SMA = IPA	Asal_Sekolah		
	3	0.150	Gender	Jurusan_SMA = IPA	Asisten		
	3	0.100	Gender	Jurusan_SMA = IPA	Rerata_SKS		
	3	0.100	Gender	Asal_Sekolah	Jurusan_SMA = IPS		
	3	0.100	Gender	Asal_Sekolah	Asisten		
	3	0.150	Gender	Asal_Sekolah	Rerata_SKS		
	3	0.050	Gender	Asal_Sekolah	Jurusan_SMA = L...		
	3	0.100	Gender	Jurusan_SMA = IPS	Rerata_SKS		

No. of Sets: 55	Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
Total Max. Size: 5	3	0.150	Gender	Asisten	Rerata_SKS		
Min. Size: 1	3	0.050	Gender	Asisten	Jurusan_SMA = L...		
Max. Size: 5	3	0.050	Gender	Rerata_SKS	Jurusan_SMA = L...		
Contains Item:	3	0.100	Jurusan_SMA = IPA	Asal_Sekolah	Asisten		
	3	0.050	Jurusan_SMA = IPA	Asal_Sekolah	Rerata_SKS		
<input type="button" value="Update View"/>	3	0.100	Jurusan_SMA = IPA	Asisten	Rerata_SKS		
	3	0.050	Asal_Sekolah	Jurusan_SMA = IPS	Rerata_SKS		
	3	0.100	Asal_Sekolah	Asisten	Rerata_SKS		
	3	0.050	Asal_Sekolah	Asisten	Jurusan_SMA = L...		
	3	0.050	Asal_Sekolah	Rerata_SKS	Jurusan_SMA = L...		
	3	0.050	Asisten	Rerata_SKS	Jurusan_SMA = L...		
	4	0.050	Gender	Jurusan_SMA = IPA	Asal_Sekolah	Asisten	
	4	0.050	Gender	Jurusan_SMA = IPA	Asal_Sekolah	Rerata_SKS	
	4	0.100	Gender	Jurusan_SMA = IPA	Asisten	Rerata_SKS	
	4	0.050	Gender	Asal_Sekolah	Jurusan_SMA = IPS	Rerata_SKS	
	4	0.100	Gender	Asal_Sekolah	Asisten	Rerata_SKS	

No. of Sets: 55	Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
Total Max. Size: 5	3	0.100	Asal_Sekolah	Asisten	Rerata_SKS		
Min. Size: 1	3	0.050	Asal_Sekolah	Asisten	Jurusan_SMA = L...		
Max. Size: 5	3	0.050	Asal_Sekolah	Rerata_SKS	Jurusan_SMA = L...		
Contains Item:	3	0.050	Asisten	Rerata_SKS	Jurusan_SMA = L...		
	4	0.050	Gender	Jurusan_SMA = IPA	Asal_Sekolah	Asisten	
	4	0.050	Gender	Jurusan_SMA = IPA	Asal_Sekolah	Rerata_SKS	
	4	0.100	Gender	Jurusan_SMA = IPA	Asisten	Rerata_SKS	
	4	0.050	Gender	Asal_Sekolah	Jurusan_SMA = IPS	Rerata_SKS	
	4	0.100	Gender	Asal_Sekolah	Asisten	Rerata_SKS	
	4	0.050	Gender	Asal_Sekolah	Asisten	Jurusan_SMA = L...	
	4	0.050	Gender	Asal_Sekolah	Rerata_SKS	Jurusan_SMA = L...	
	4	0.050	Gender	Asisten	Rerata_SKS	Jurusan_SMA = L...	
	4	0.050	Jurusan_SMA = IPA	Asal_Sekolah	Asisten	Rerata_SKS	
	4	0.050	Asal_Sekolah	Asisten	Rerata_SKS	Jurusan_SMA = L...	
	5	0.050	Gender	Jurusan_SMA = IPA	Asal_Sekolah	Asisten	Rerata_SKS
	5	0.050	Gender	Asal_Sekolah	Asisten	Rerata_SKS	Jurusan_SMA = L...

Show rules matching
all of these conclusions:

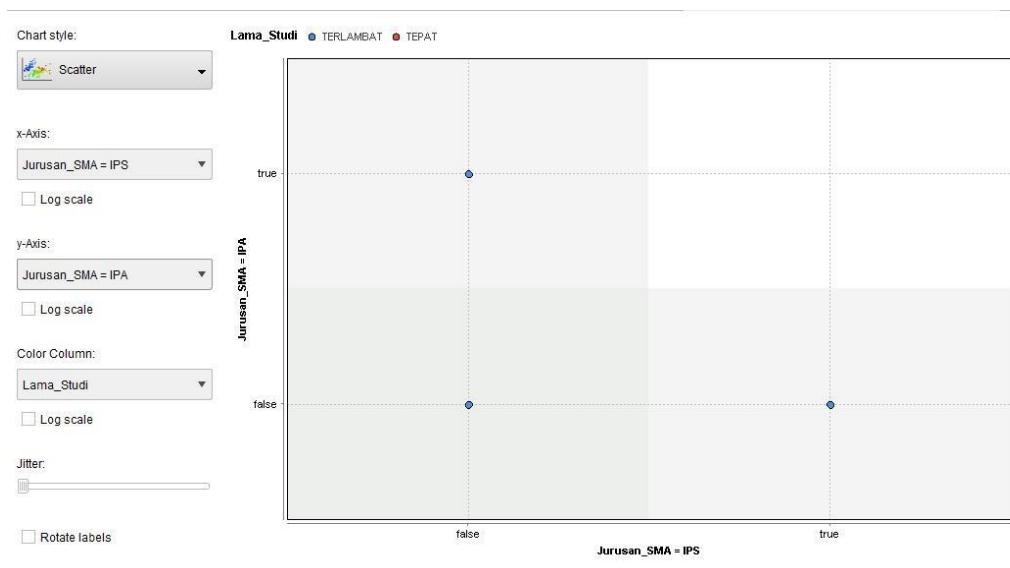
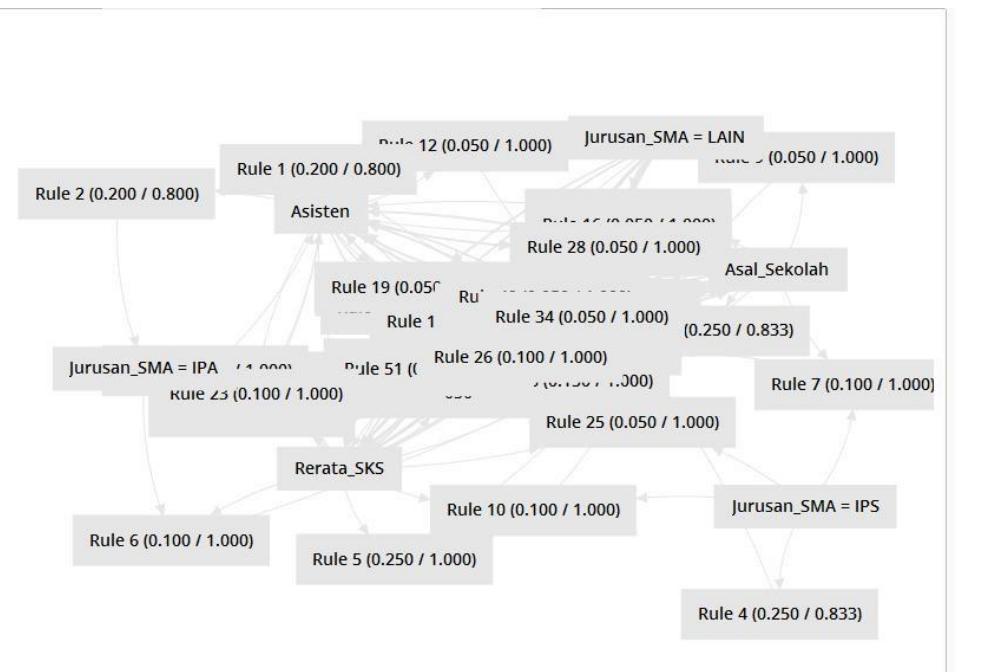
Gender
Jurusan_SMA = IPA
Asal_Sekolah
Asisten
Rerata_SKS

Min. Criterion:
confidence

Min. Criterion Value:

No.	Premises	Conclusion	Support	Confidence
3	Asal_Sekolah	Gender	0.250	0.8
4	Jurusan_SMA = IPS	Gender	0.250	0.8
5	Rerata_SKS	Gender	0.250	1
6	Jurusan_SMA = IPA, Rerata_SKS	Gender	0.100	1
7	Asal_Sekolah, Jurusan_SMA = IPS	Gender	0.100	1
8	Asal_Sekolah, Rerata_SKS	Gender	0.150	1
9	Asal_Sekolah, Jurusan_SMA = LAIN	Gender	0.050	1
10	Jurusan_SMA = IPS, Rerata_SKS	Gender	0.100	1
11	Asisten, Rerata_SKS	Gender	0.150	1
12	Asisten, Jurusan_SMA = LAIN	Gender	0.050	1
13	Rerata_SKS, Jurusan_SMA = LAIN	Gender	0.050	1
14	Jurusan_SMA = IPA, Rerata_SKS	Asisten	0.100	1
15	Asal_Sekolah, Jurusan_SMA = LAIN	Asisten	0.050	1
16	Asisten, Jurusan_SMA = LAIN	Asal_Sekolah	0.050	1
17	Asal_Sekolah, Jurusan_SMA = LAIN	Rerata_SKS	0.050	1
18	Rerata_SKS, Jurusan_SMA = LAIN	Asal_Sekolah	0.050	1

	Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
	0.250	0.833	0.962	-0.350	0.025	1.111	1.500
	0.250	0.833	0.962	-0.350	0.025	1.111	1.500
	0.250	1	1	-0.250	0.062	1.333	∞
	0.100	1	1	-0.100	0.025	1.333	∞
	0.100	1	1	-0.100	0.025	1.333	∞
	0.150	1	1	-0.150	0.038	1.333	∞
	0.050	1	1	-0.050	0.012	1.333	∞
	0.100	1	1	-0.100	0.025	1.333	∞
	0.150	1	1	-0.150	0.038	1.333	∞
	0.050	1	1	-0.050	0.012	1.333	∞
	0.050	1	1	-0.050	0.012	1.333	∞
	0.100	1	1	-0.100	0.075	4	∞
	0.050	1	1	-0.050	0.038	4	∞
	0.050	1	1	-0.050	0.035	3.333	∞
	0.050	1	1	-0.050	0.038	4	∞
	0.050	1	1	-0.050	0.035	3.333	∞



MODUL 12
(Regresi Linier Sederhana)
KEGIATAN

Kegiatan 1

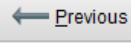
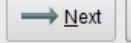
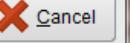
Import Data - Format your columns.

Format your columns.

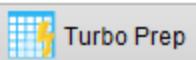
Replace errors with missing values ⓘ

NO_SISWA polynomial id	NAMA polynomial	LAMA BELAJAR (JAM) integer	NILAI integer
1 S-101	JOKO	15	783
2 S-102	AGUS	18	877
3 S-103	SUSI	7	505
4 S-104	DYAH	9	860
5 S-105	WATI	15	968
6 S-106	IKA	17	793
7 S-107	EKO	10	752
8 S-108	YANTO	5	571
9 S-109	WAWAN	8	667
10 S-110	MAHMUD	15	723

 no problems.

Open in

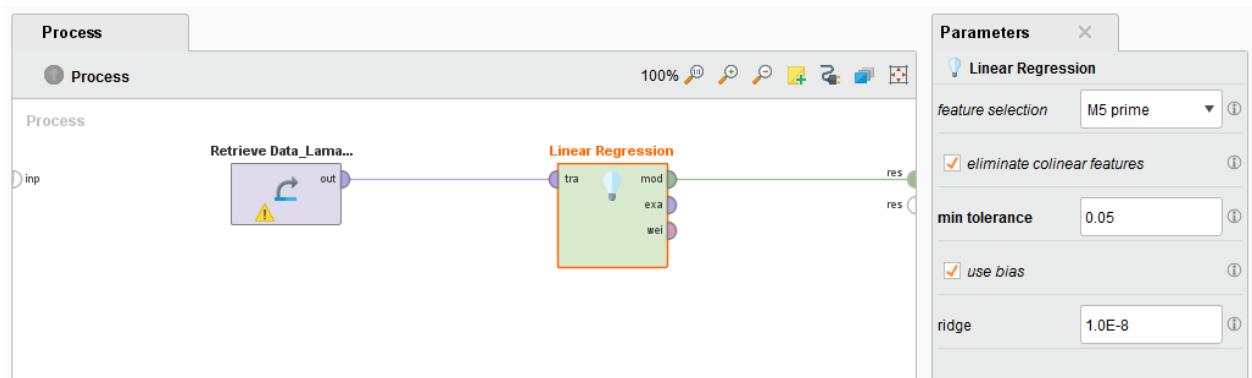


Turbo Prep



Auto Model

Row No.	NO_SISWA	NILAI	LAMA BELA...
1	S-101	783	15
2	S-102	877	18
3	S-103	505	7
4	S-104	860	9
5	S-105	968	15
6	S-106	793	17
7	S-107	752	10
8	S-108	571	5
9	S-109	667	8
10	S-110	723	15



Attribute	Coefficient	Std. Error	Std. Coeffici...	Tolerance	t-Stat	p-Value	Code
LAMA BELAJA...	21.608	7.645	0.707	1	2.827	0.022	**
(Intercept)	492.769	96.909	?	?	5.085	0.001	****

Linear Regression

$$21.608 * \text{LAMA BELAJAR (JAM)} + 492.769$$

NO_SISWA	NAMA	LAMA BELAJAR (JAM)	NILAI	Y
S-101	JOKO	15	783	816.889
S-102	AGUS	18	877	881.713
S-103	SUSI	7	505	644.025
S-104	DYAH	9	860	687.241
S-105	WATI	15	968	816.889
S-106	IKA	17	793	860.105
S-107	EKO	10	752	708.849
S-108	YANTO	5	571	600.809
S-109	WAWAN	8	667	665.633
S-110	MAHMUD	15	723	816.889

Kegiatan 2

Import Data - Format your columns.

Format your columns.

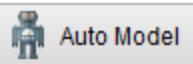
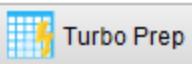
Replace errors with missing values ⓘ

NO_SISWA <small>polynominal id</small>	NAMA <small>polynominal</small>	LAMA BELAJAR (JAM) <small>integer</small>
1 S-111	BUDI	12
2 S-112	SANTI	13
3 S-113	DIAN	14
4 S-114	DANI	11
5 S-115	AHMAD	5
6 S-116	BAYU	13
7 S-117	RISA	9
8 S-118	RANI	10
9 S-119	YANI	10
10 S-120	RATIH	9

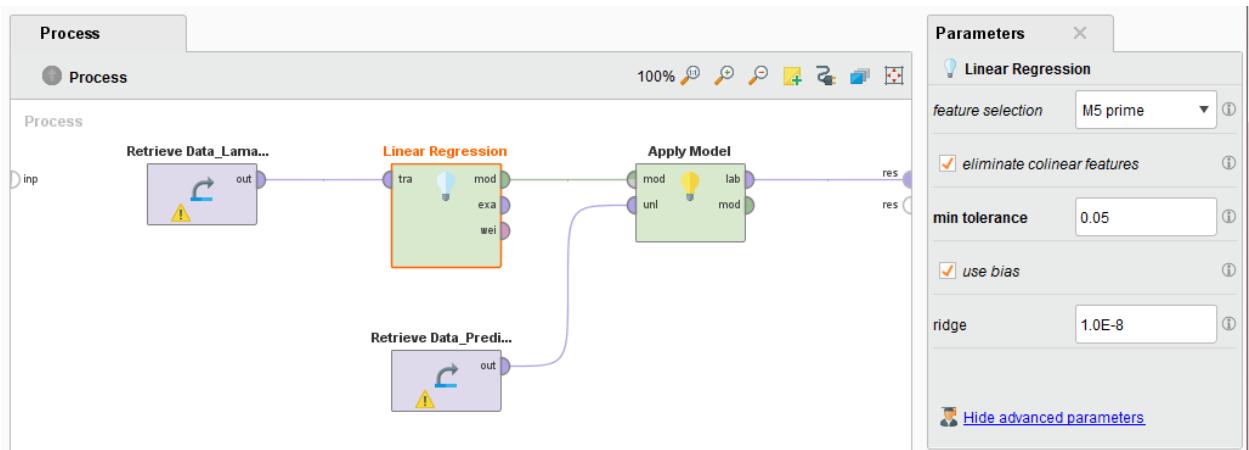
 no problems.

 [Previous](#)  [Next](#)  [Cancel](#)

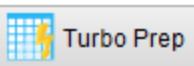
Open in



Row No.	NO_SISWA	LAMA BELA...
1	S-111	12
2	S-112	13
3	S-113	14
4	S-114	11
5	S-115	5
6	S-116	13
7	S-117	9
8	S-118	10
9	S-119	10
10	S-120	9



Open in

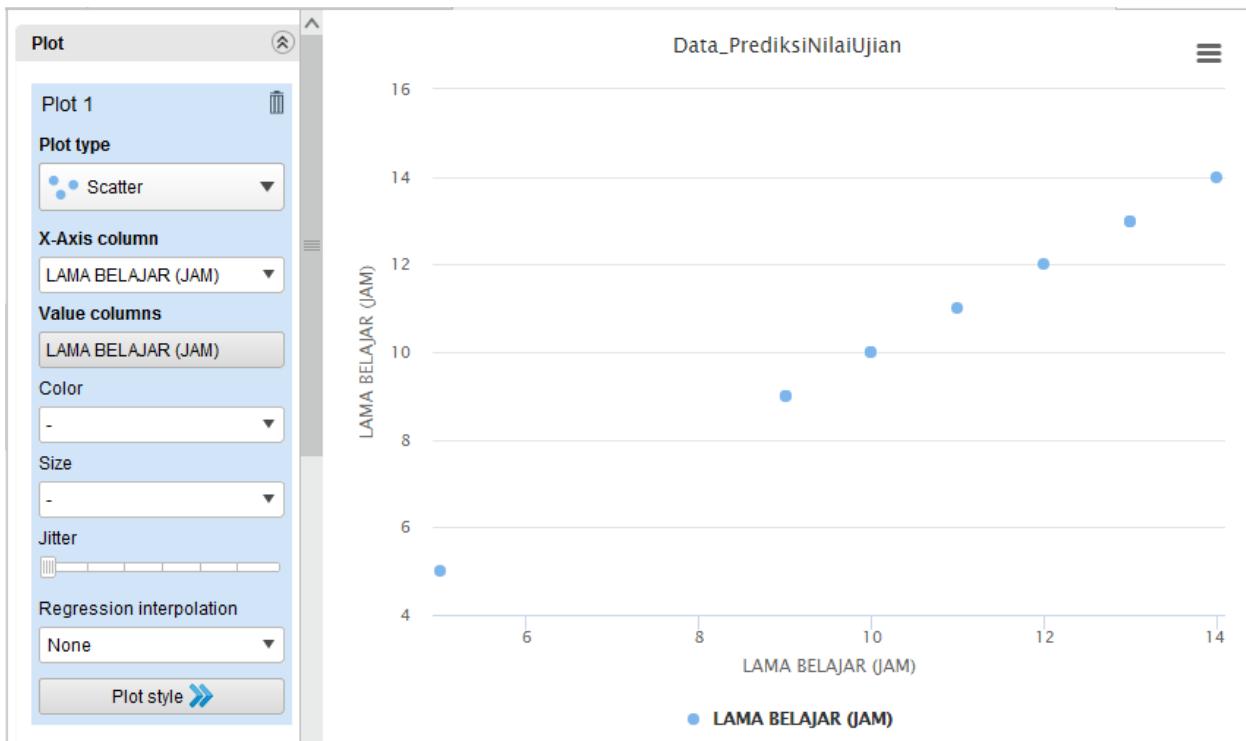


Turbo Prep



Auto Model

Row No.	NO_SISWA	prediction(No...)	LAMA BELA...
1	S-111	752.061	12
2	S-112	773.668	13
3	S-113	795.276	14
4	S-114	730.453	11
5	S-115	600.807	5
6	S-116	773.668	13
7	S-117	687.238	9
8	S-118	708.845	10
9	S-119	708.845	10
10	S-120	687.238	9



NO_SISWA	NAMA	LAMA BELAJAR (JAM)	Tabel	Model Regresi
S-111	BUDI	12	752.061	752.065
S-112	SANTI	13	773.668	773.673
S-113	DIAN	14	795.276	795.281
S-114	DANI	11	730.453	730.457
S-115	AHMAD	5	600.807	600.809
S-116	BAYU	13	773.668	773.673
S-117	RISA	9	687.238	687.241
S-118	RANI	10	708.845	708.849
S-119	YANI	10	708.845	708.849
S-120	RATIH	9	687.238	687.241

MODUL 12
(Regresi Linier Sederhana)
TUGAS

Tugas 1

Import Data - Format your columns.

Format your columns.

Replace errors with missing values (i)

NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA K...	DAYA BELI (RUPIAH)
1	1000000	6	834000
2	1400000	7	1200000
3	200000	3	134000
4	1400000	6	1167000
5	500000	3	334000
6	1700000	5	1360000
7	400000	3	267000
8	1900000	5	1520000
9	300000	3	200000
10	500000	4	375000
11	700000	7	600000
12	1900000	3	1267000
13	800000	4	600000

no problems.

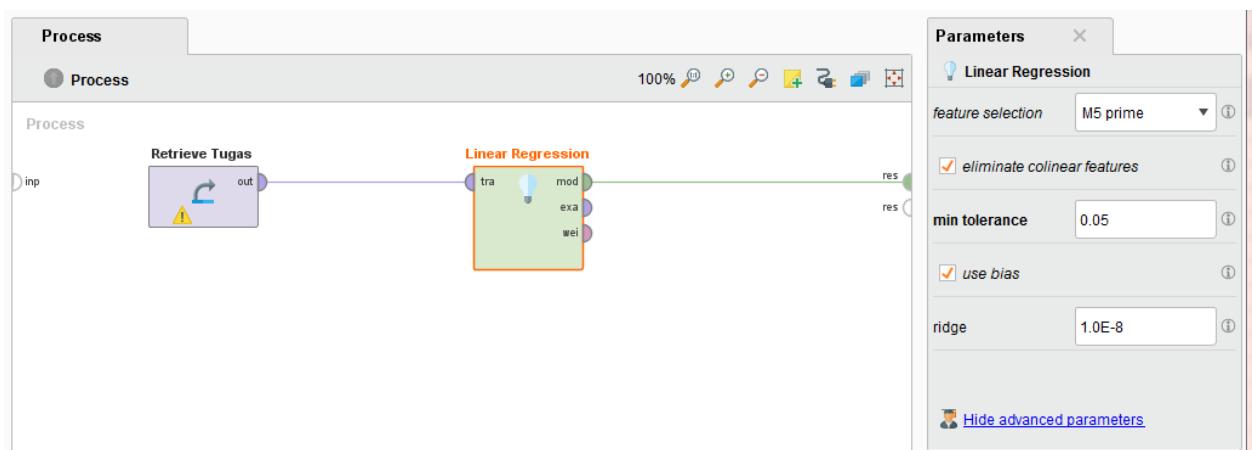
Open in

Turbo Prep

Auto Model

Row No.	NO. RESPON...	DAYA BELI (...)	PENDAPATA...	JUMLAH AN...
1	1	834000	1000000	6
2	2	1200000	1400000	7
3	3	134000	200000	3
4	4	1167000	1400000	6
5	5	334000	500000	3
6	6	1360000	1700000	5
7	7	267000	400000	3
8	8	1520000	1900000	5
9	9	200000	300000	3
10	10	375000	500000	4
11	11	600000	700000	7
12	12	1267000	1900000	3
13	13	600000	800000	4

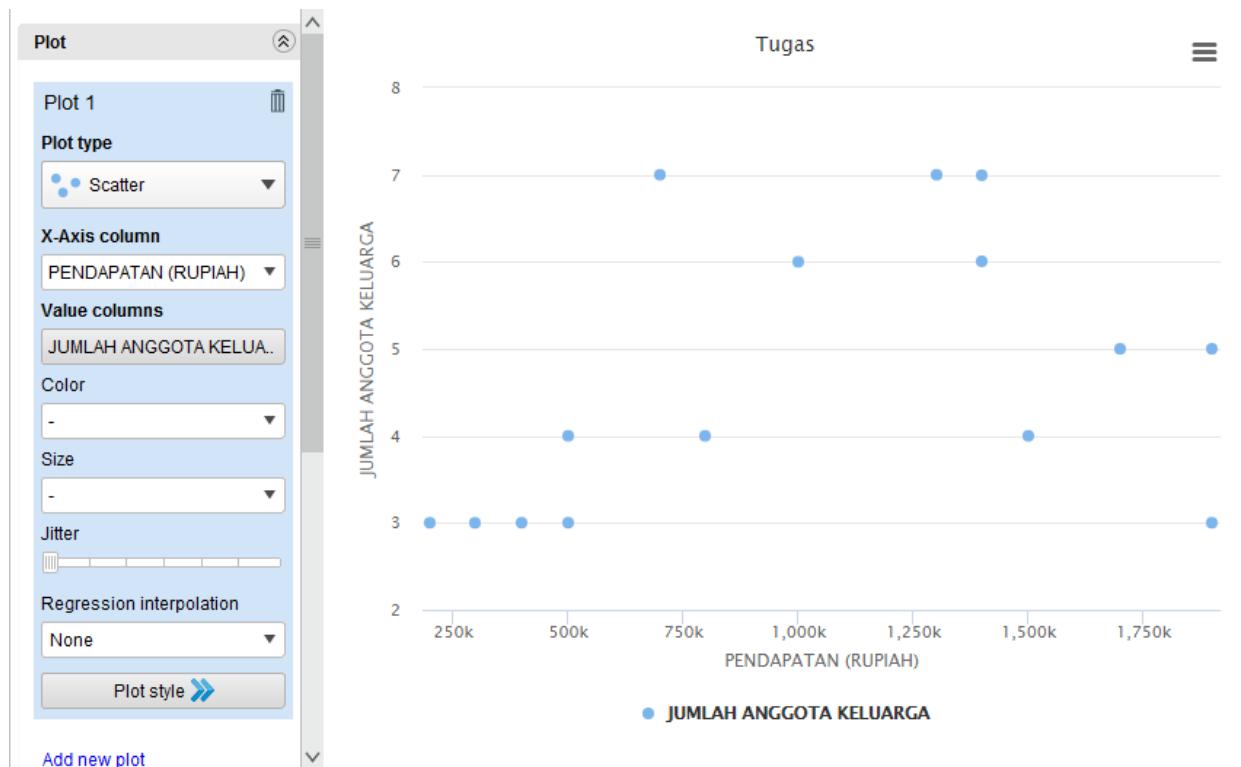
ExampleSet (15 examples, 2 special attributes, 2 regular attributes)



Attribute	Coefficient	Std. Error	Std. Coeffici...	Tolerance	t-Stat	p-Value	Code
PENDAPATAN...	0.739	0.021	0.924	0.857	35.037	0.000	****
JUMLAH ANG...	47807.624	7833.319	0.161	0.857	6.103	0.000	****
(Intercept)	-180222.487	36497.284	?	?	-4.938	0.000	****

LinearRegression

```
0.739 * PENDAPATAN (RUPIAH)
+ 47807.624 * JUMLAH ANGGOTA KELUARGA
- 180222.487
```



Tugas 2

Import Data - Format your columns.

Format your columns.

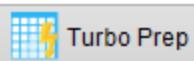
Replace errors with missing values (i)

	NO. RESPONDEN <small>integer id</small>	PENDAPATAN (RUPIAH) <small>integer</small>	JUMLAH ANGGOTA KELUARGA <small>integer label</small>
1	1	900000	5
2	2	800000	3
3	3	500000	2
4	4	1900000	6
5	5	600000	2
6	6	800000	5
7	7	1000000	6
8	8	1100000	4
9	9	1000000	4
10	10	500000	3

 no problems.

 [Previous](#)  [Next](#)  [Cancel](#)

Open in

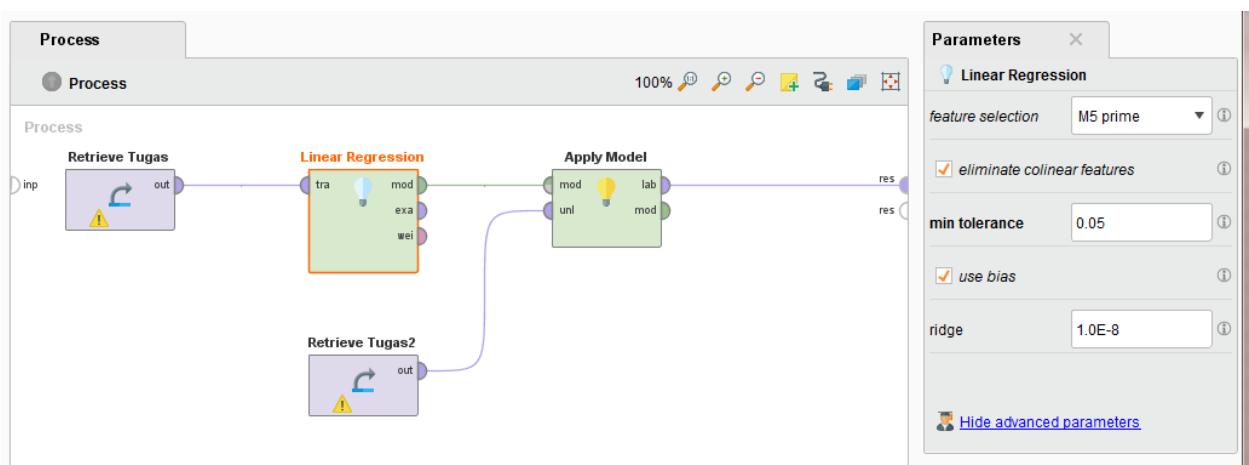


Turbo Prep



Auto Model

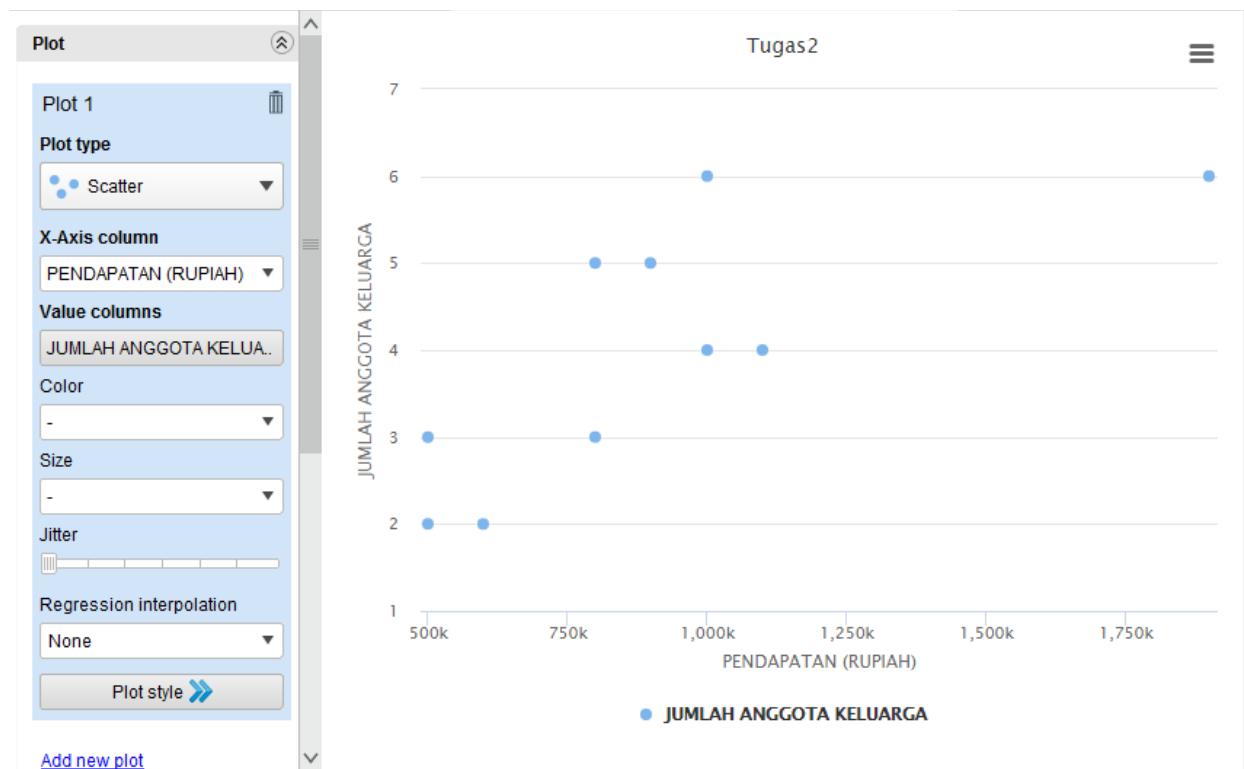
Row No.	NO. RESPON...	JUMLAH AN...	PENDAPATA...
1	1	5	900000
2	2	3	800000
3	3	2	500000
4	4	6	1900000
5	5	2	600000
6	6	5	800000
7	7	6	1000000
8	8	4	1100000
9	9	4	1000000
10	10	3	500000



Open in  Turbo Prep

 Auto Model

Row No.	NO. RESPON...	JUMLAH AN...	prediction(D...	PENDAPATA...
1	1	5	723933.263	900000
2	2	3	554416.056	800000
3	3	2	284902.556	500000
4	4	6	1510760.476	1900000
5	5	2	358804.515	600000
6	6	5	650031.304	800000
7	7	6	845642.845	1000000
8	8	4	823929.557	1100000
9	9	4	750027.598	1000000
10	10	3	332710.179	500000



NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	Y
1	900000	5	723915.633
2	800000	3	554400.385
3	500000	2	284892.761
4	1900000	6	1510723.257
5	600000	2	358792.761
6	800000	5	650015.633
7	1000000	6	845623.257
8	1100000	4	823908.009
9	1000000	4	750008.009
10	500000	3	332700.385