

REKAP TUGAS
Praktikum Data Warehousing dan Data Mining



Disusun Oleh :

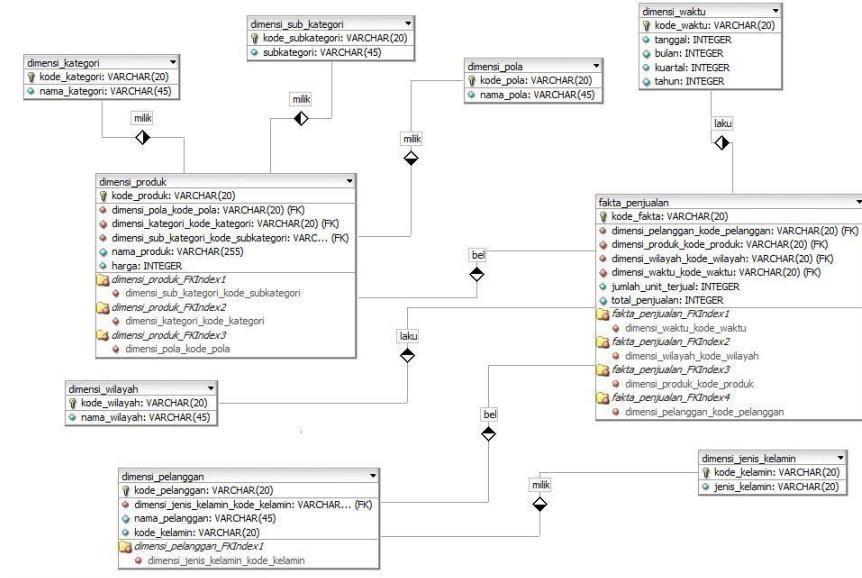
Nama : Cinde Prawito P

NIM : L200170120

Kelas : E

PROGRAM STUDI INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA

MODUL 1



MODUL 5

Kegiatan

- Membuat tabel pada excel disimpan dengan nama “**Fakta_Penjualan.xls**”

The screenshot shows a Microsoft Excel spreadsheet titled "Sheet1". The table structure is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	bulan	kuartal	tahun	nama_produk	nama_kategori	nama_subkategori	nama_pola	nama_pelanggan	jenis_kelamin	nama_wilayah	jumlah	harga			
2	12	4	2011	Jarik Standar Print Sogan	Standar	Jarik	Print	Bapak Ketut	PRIA	Bali	2	225000			
3	1	1	2012	Kaos Batik Cap Lukis	Batik	Kaos	Cap	Ibu Harini	WANITA	Jawa Timur	14	30000			
4	4	2	2012	Jarik Standar Tulis Sarimbit	Standar	Jarik	Tulis	Ibu Harini	WANITA	Jawa Timur	4	40000			
5	4	2	2011	Hem Katun Print Kawung	Katun	Hem	Print	Ibu Harini	WANITA	Jawa Timur	3	70000			
6	9	3	2012	Batik Standar Cap Tumpal	Standar	Batik	Cap	Bapak Heru	PRIA	Jawa Timur	1	150000			
7	5	2	2012	Hem Katun Print Kelengka	Katun	Hem	Print	Bapak Totok	PRIA	Jawa Timur	3	299000			
8	12	4	2011	Balero Standar Cap Sido Mukti	Standar	Balero	Cap	Ibu Hatamah	WANITA	Jawa Timur	1	225000			
9	10	4	2011	Sarimbit Standar Print Lukis	Standar	sarimbit	Print	Ibu Hatamah	WANITA	Jawa Timur	1	150000			
10	1	1	2011	Kaos Katun Print Bola	Katun	Kaos	Print	Bapak Imron	PRIA	Jawa Barat	1	60000			
11	2	1	2012	Celana Standar Cap Warna	Standar	Celana	Cap	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000			
12	3	1	2010	Celana Standar cap Warna	Standar	Celana	Print	Ibu Hadi Sukarni	WANITA	Jawa Barat	17	55000			
13	3	1	2011	Bahan Standar Cap Lansen	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	3	2012	Jam Standar Print Lukis	Standar	Jam	Print	Ibu Siti Arya	WANITA	Jawa Barat	44	80000			
16	9	3	2012	Hem Standar Cap Tumpal	Standar	Hem	Cap	Ibu Aini Kasmaji	WANITA	Jawa Tengah	1	100000			
17	6	2	2012	Bahan Lawasan Tulis Toilet	Lawasan	Bahan	Tulis	Ibu Niken	WANITA	Jawa Tengah	1	130000			
18	8	3	2011	Hem Standar Tulis Madura	Standar	Hem	Tulis	Ibu Atik	WANITA	Jawa Tengah	5	55000			
19	4	2	2012	Bahan Standar Cap Garis	Standar	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	7	135000			
20	6	2	2010	Beludru Cap Mahkota	Beludru	Bahan	Cap	Ibu Tyas	WANITA	Jawa Tengah	1	500000			
21	11	4	2010	Hem Sutra Print Rama	Sutra	Hem	Print	Ibu Tyas	WANITA	Jawa Tengah	5	100000			
22															
23															
24															

- Pada Pivot Table Field List, menyusun layout field dengan urutan :

- Field **nama_subkategori** di kotak Row Labels
- Field **tahun** di kotak Column Labels
- Field **jumlah** di kotak Values

Dan akan menghasilkan suatu tabel dengan grouping field **nama_subkategori** pada bagian baris, field **tahun** pada kolom. Sedangkan nilai total **jumlah** ditempatkan pada cell-cell hasil perpotongan item grouping baris dan kolom tersebut.

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable. The table structure is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
3	Sum of jumlah	Column Labels													
4	Row Labels		2010	2011	2012	Grand Total									
5	Bahan		1	8	8	17									
6	Balero			1		1									
7	Batik				1	1									
8	Celana		17		17	34									
9	Hem			5	8	4	17								
10	Jam				2	4	44	44							
11	Jarik				1	14	15								
12	Kaos					1	1	6							
13	Rok						1								
14	sarimbit							1							
15	Grand Total		23	21	93	137									

The PivotTable Fields dialog box is open on the right side of the screen, showing the following settings:

- Choose fields to add to report: jumlah
- Drag fields between areas below:
 - Filters: tahun
 - Columns: nama_subkategori
 - Rows: nama_subkategori
 - Values: Sum of jumlah

- Kemudian menambahkan field jumlah pada kotak values dan mensetting jumlah2 di values field setting dengan mengubah nilai sum menjadi count dan akan menghasilkan tabel seperti yang di bawah ini :

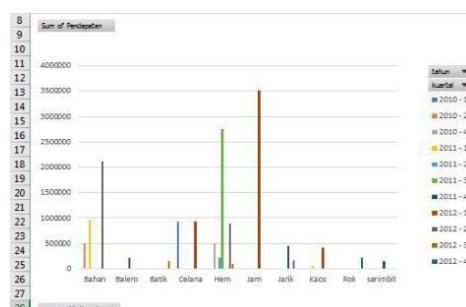
SEBELUM DIUBAH

SESUDAH DIUBAH

- Calculated Field, jika kita ingin mencari jumlah pendapatan yang diperoleh berdasarkan jumlah produk yang terjual dikalikan dengan harga produk menggunakan Pivot Table dengan menggunakan **Calculated Field**

- Operasi Roll Up dan Drill Down, digunakan untuk melihat data secara lebih rinci dan secara lebih umum berdasarkan kategori tertentu pada sebuah data, seperti gambar dibawah ini

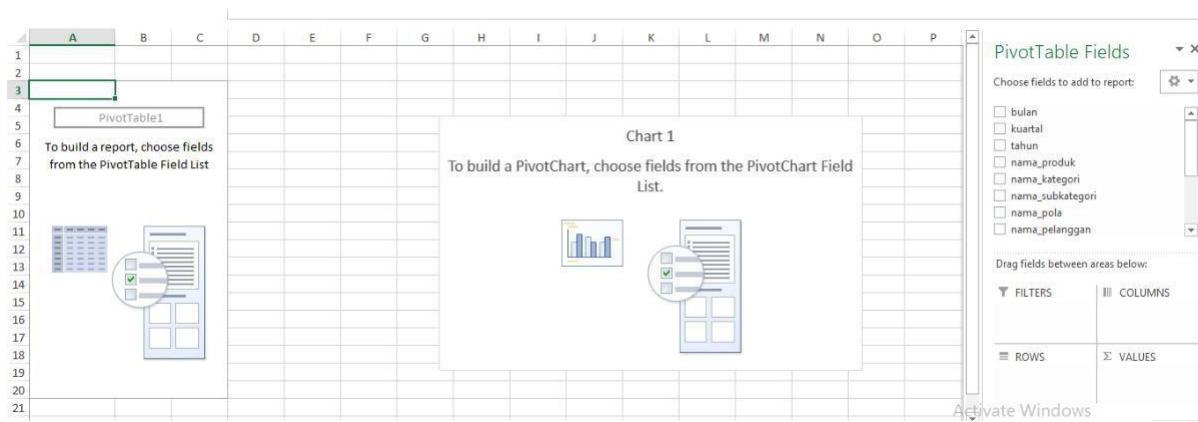
- Gambar dibawah menggunakan PivotChart yang merupakan sebuah cara untuk menampilkan cube dalam bentuk grafik.



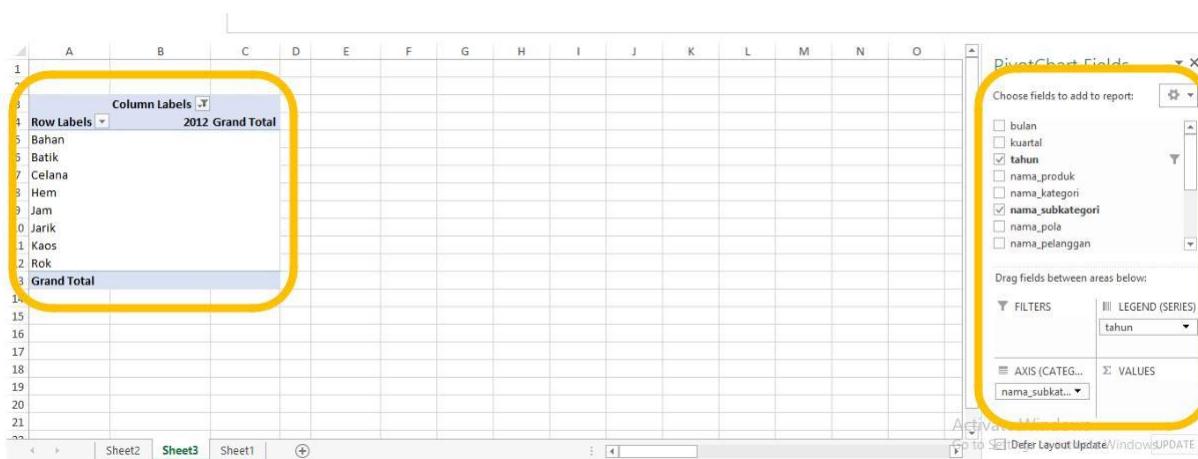
Tugas

NO 1

- Membuat PivotTable dan PivotChart baru



- Menambahkan field **nama_subkategori** ke kotak Row Labels dan field **tahun** ke kotak Column Labels dan tahun yang di gunakan hanya tahun 2012



- Menambahkan field Pendapatan pada kotak values

PivotTable Fields

Choose fields to add to report:

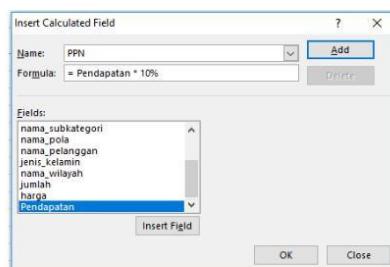
- `nama_pola`
- `nama_pelanggan`
- `jenis_kelamin`
- `nama_wilayah`
- `jumlah`
- `harga`
- Pendapatan**

Drag fields between areas below:

FILTERS **LEGEND (SERIES)**
 tahun

ROWS **COLUMNS**
 nama_subkata... Σ VALUES
 Sum of Pendapatan

- Kemudian tambahkan field baru dengan cara “Insert Calculated Field” dengan nama PPN dan formulanya = Pendapatan * 10%. Dan klik OK



PivotTable Fields

Choose fields to add to report:

- `nama_pola`
- `nama_pelanggan`
- `jenis_kelamin`
- `nama_wilayah`
- `jumlah`
- Pendapatan**
- PPN**

Drag fields between areas below:

FILTERS **COLUMNS**
 tahun Σ Values

ROWS **VALUES**
 nama_subkata... Σ Values
 Sum of Pendapatan Σ Sum of PPN

Kemudian tambahkan field baru dengan cara “Insert Calculated Field” dengan nama **Total Penghasilan** dan formulanya = Pendapatan – PPN. Dan klik OK

Insert Calculated Field

Name: Total Penghasilan
Formula: = Pendapatan - PPN

Fields:
nama_pola
nama_pelanggan
jenis_kelamin
nama_wilayah
jumlah
harga
Pendapatan
PPN

OK Close

	A	B	C	D	E	F	G	H	I
1									
2									
3		Column Labels							
4		2012			Total Sum of Pendapatan	Total Sum of PPN	Total Sum of Total Penghasilan		
5	Row Labels	Sum of Pendapatan	Sum of PPN	Sum of Total Penghasilan					
6	Bahan	2120000	212000	1908000	2120000	212000	1908000		
7	Batik	150000	15000	135000	150000	15000	135000		
8	Celana	935000	93500	841500	935000	93500	841500		
9	Hem	1596000	159600	1436400	1596000	159600	1436400		
10	Jam	3520000	352000	3168000	3520000	352000	3168000		
11	Jarik	160000	16000	144000	160000	16000	144000		
12	Kaos	420000	42000	378000	420000	42000	378000		
13	Rok	225000	22500	202500	225000	22500	202500		
14	Grand Total	115692000	11569200	104122800	115692000	11569200	104122800		
15									
16									
17									
18									
19									
20									
21									

Sheet2 Sheet3 Sheet1 +

PivotTable Fields

Choose fields to add to report:

- nama_pola
- nama_pelanggan
- jenis_kelamin
- nama_wilayah
- jumlah
- harga
- Pendapatan
- PPN
- Total Penghasilan

Drag fields between areas below:

FILTERS

COLUMNS

ROWS

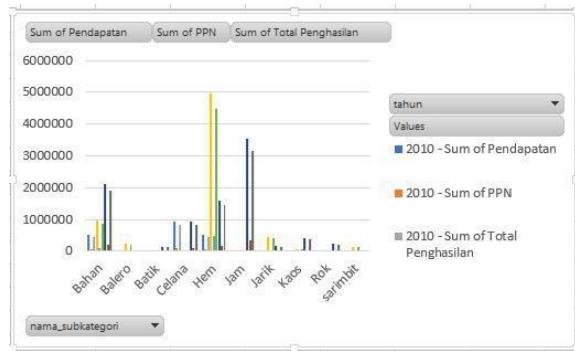
VALUES

Activate Windows

Go to: Defer Layout Update/WindowsUPDATE

NO 2

- PivotChart dari tahun 2010-2012



- PivotTable dari tahun 2010-2012

The PivotTable shows sales data for 2010, 2011, and 2012 across categories like Bahan, Batik, Celana, Hem, Jam, Jarik, Kaos, Rok, and sarimbit. The report includes columns for Sum of Pendapatan, Sum of PPN, and Sum of Total Penghasilan.

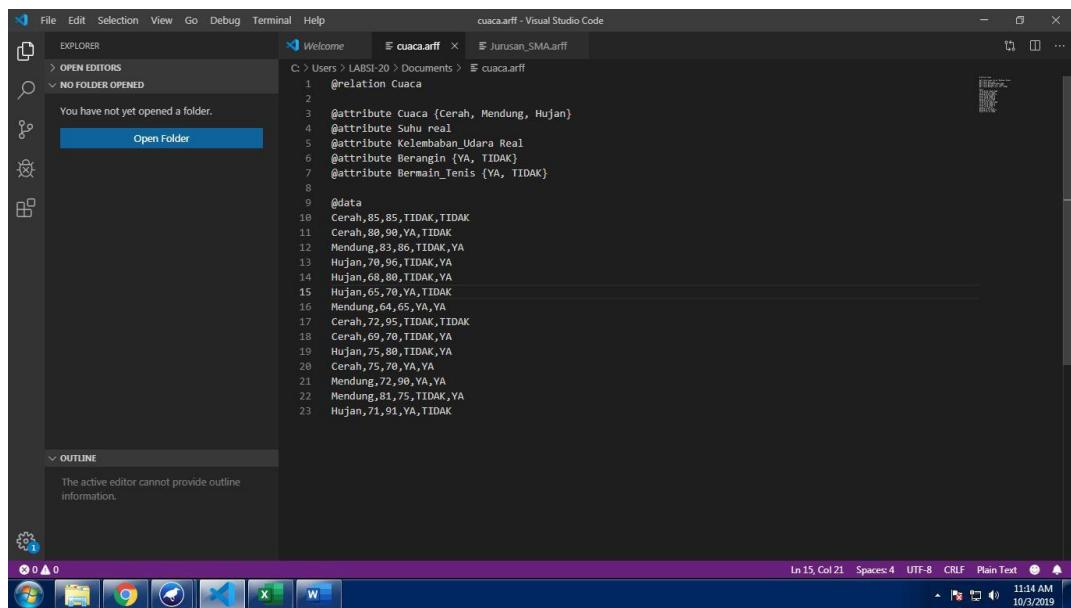
Row Labels	2010			2011			2012		
	Sum of Pendapatan	Sum of PPN	Sum of Total Penghasilan	Sum of Pendapatan	Sum of PPN	Sum of Total Penghasilan	Sum of Pendapatan	Sum of PPN	Sum of Total Penghasilan
Bahan	500000	50000	450000	960000	96000	864000	2120000	212000	190000
Balero	0	0	0	225000	22500	202500	0	0	0
Batik	0	0	0	0	0	0	150000	15000	15000
Celana	935000	93500	841500	0	0	0	935000	93500	841500
Hem	500000	50000	450000	4960000	496000	4464000	1596000	159600	145600
Jam	0	0	0	0	0	0	3520000	352000	316000
Jarik	0	0	0	450000	45000	405000	1600000	160000	140000
Kaos	0	0	0	60000	6000	54000	420000	42000	37000
Rok	0	0	0	0	0	0	225000	22500	20000
sarimbit	0	0	0	150000	15000	135000	0	0	0
Grand Total	15065000	1506500	13558500	29400000	2940000	26460000	115692000	11569200	1041200

MODUL 6

MODUL 7

Kegiatan

1. Cuaca.arff



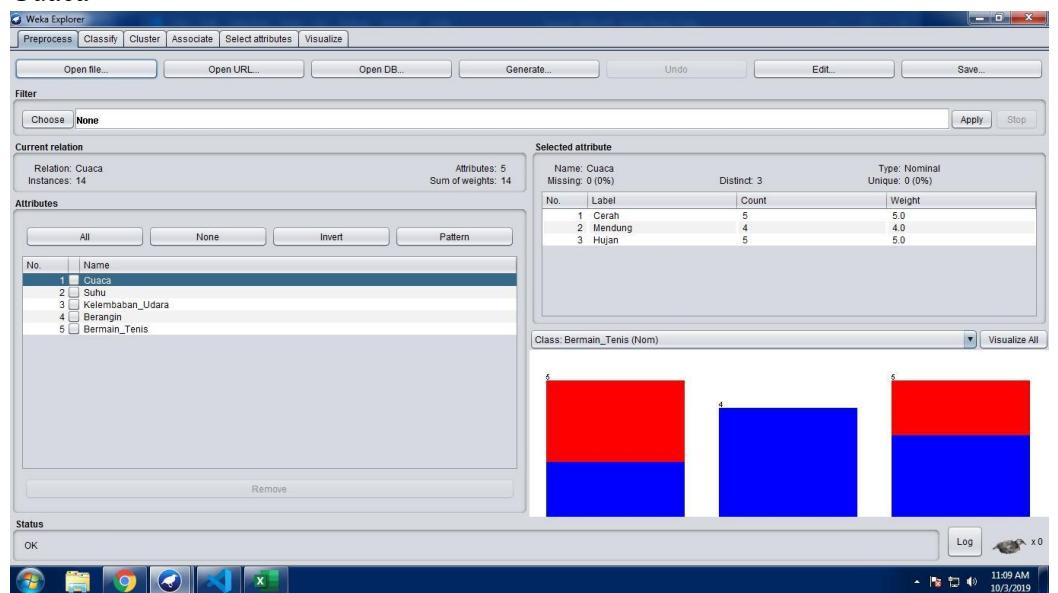
The screenshot shows the Visual Studio Code interface with the file "cuaca.arff" open. The code editor displays the following ARFF file content:

```
@relation Cuaca
@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_Udara Real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

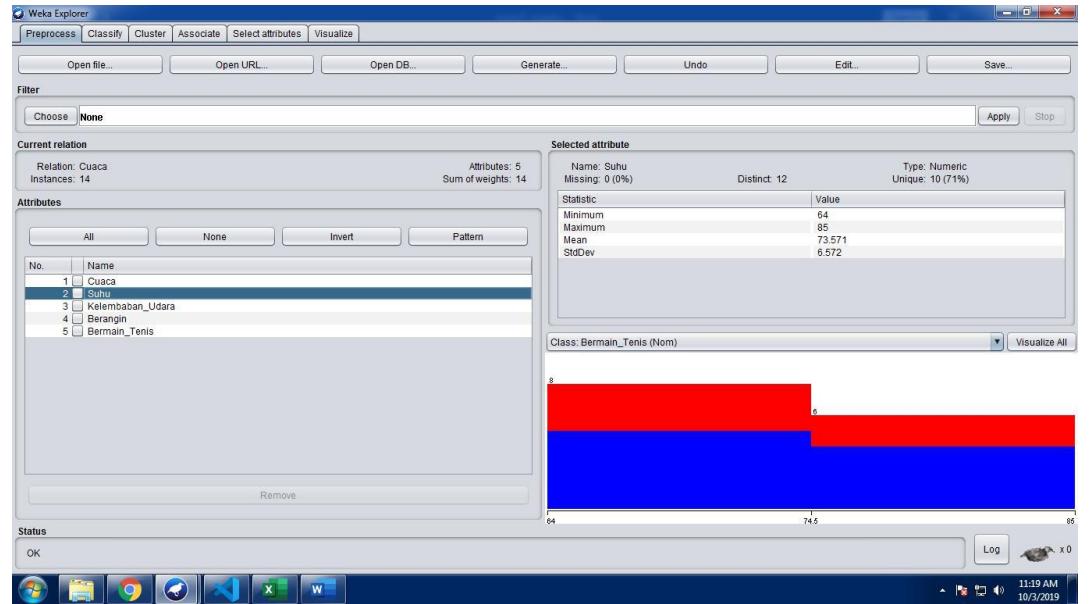
@data
Cerah, 85, 85, TIDAK, TIDAK
Cerah, 88, 98, YA, TIDAK
Mendung, 83, 86, TIDAK, YA
Hujan, 78, 96, TIDAK, YA
Hujan, 68, 88, TIDAK, YA
Hujan, 65, 78, YA, TIDAK
Mendung, 64, 65, YA, YA
Cerah, 72, 95, TIDAK, TIDAK
Cerah, 69, 70, TIDAK, YA
Hujan, 75, 80, TIDAK, YA
Cerah, 75, 78, YA, YA
Mendung, 72, 98, YA, YA
Mendung, 81, 75, TIDAK, YA
Hujan, 71, 91, YA, TIDAK
```

2. Gambar grafik dari WEKA

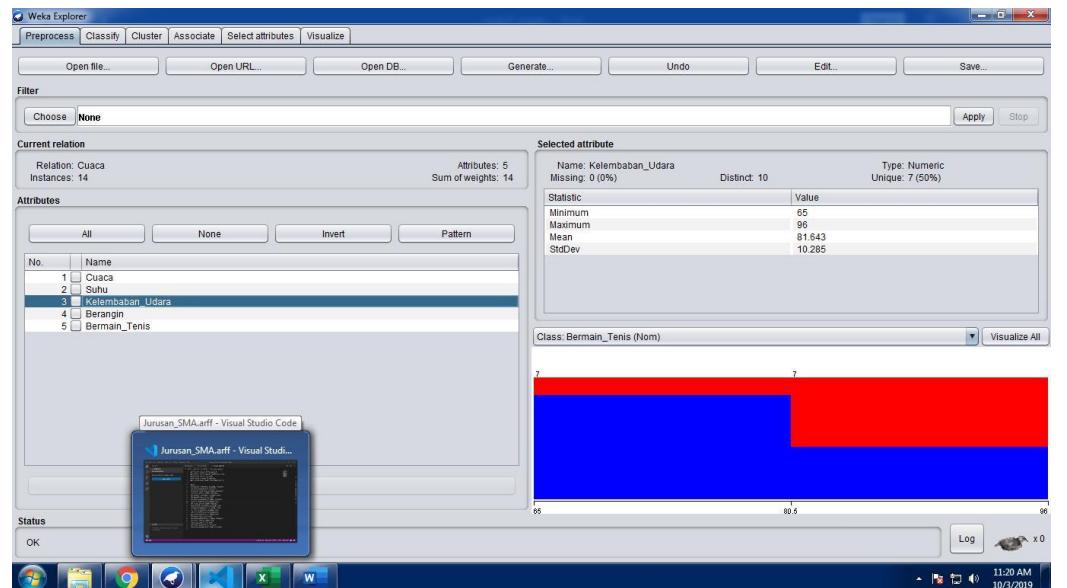
➤ Cuaca



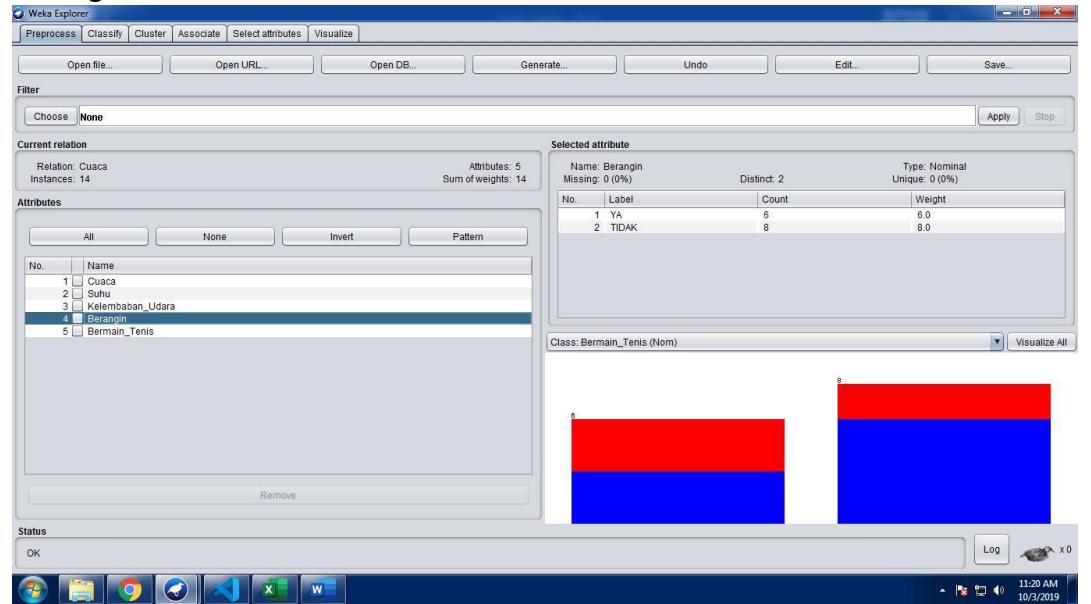
➤ Suhu



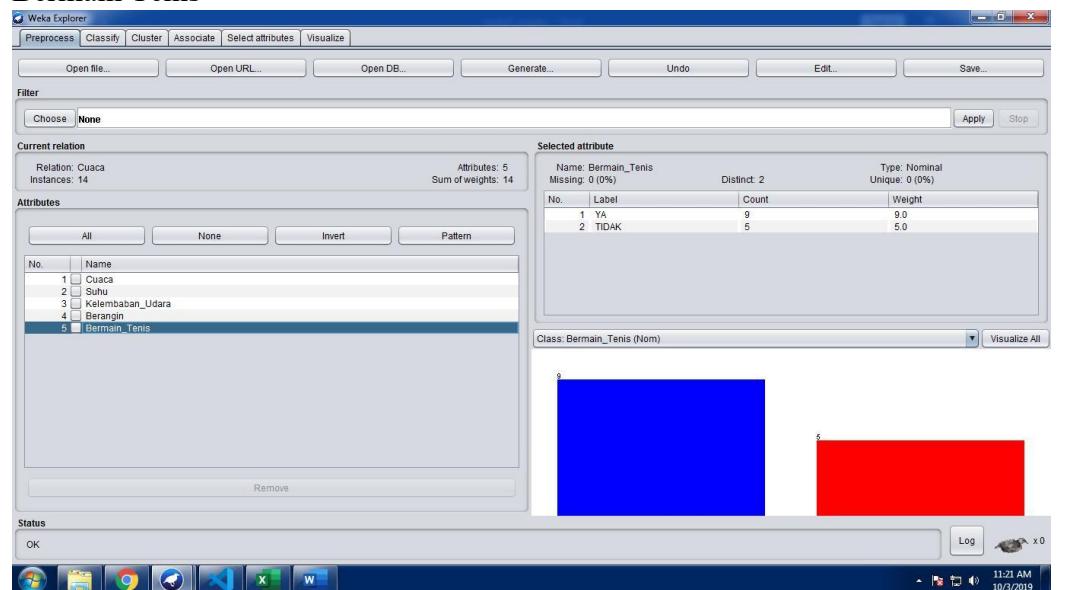
➤ Kelembaban Udara



➤ Berangin

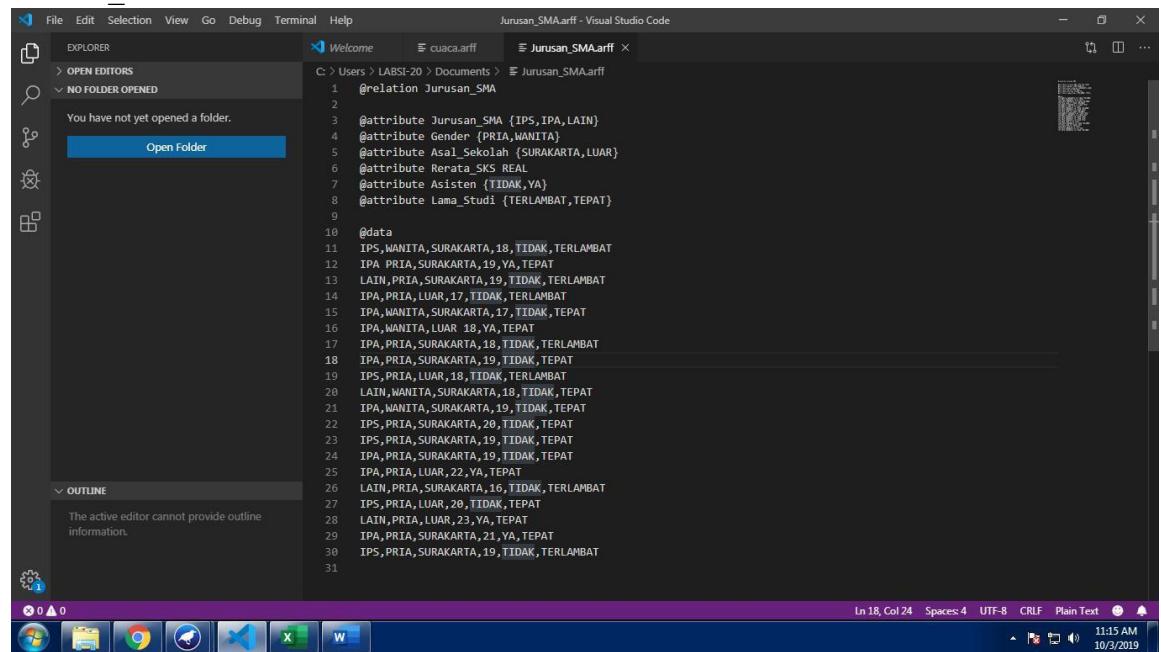


➤ Bermain Tenis



Tugas

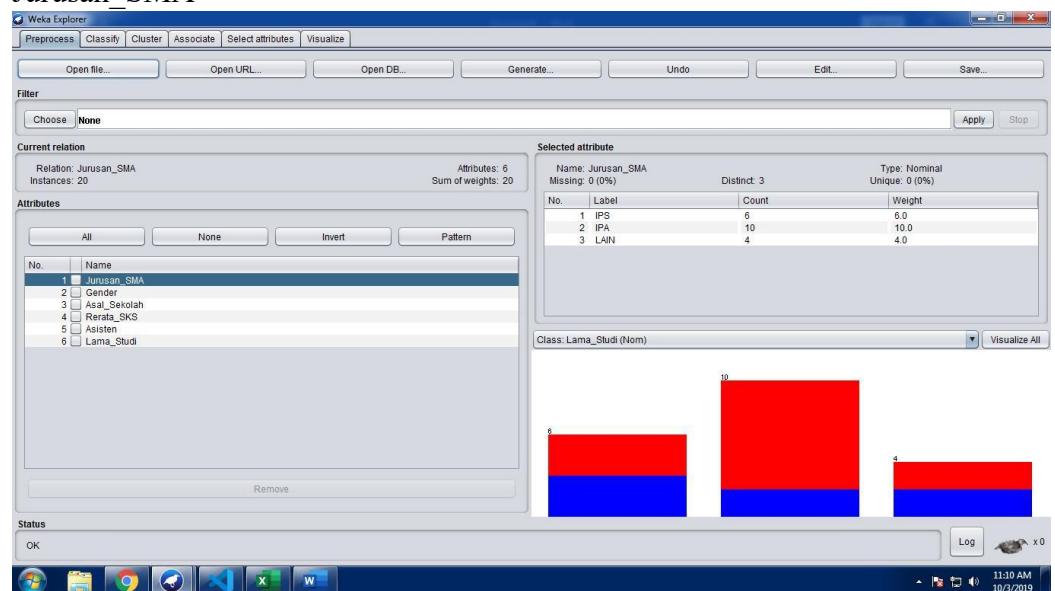
1. Jurusan_SMA.arff



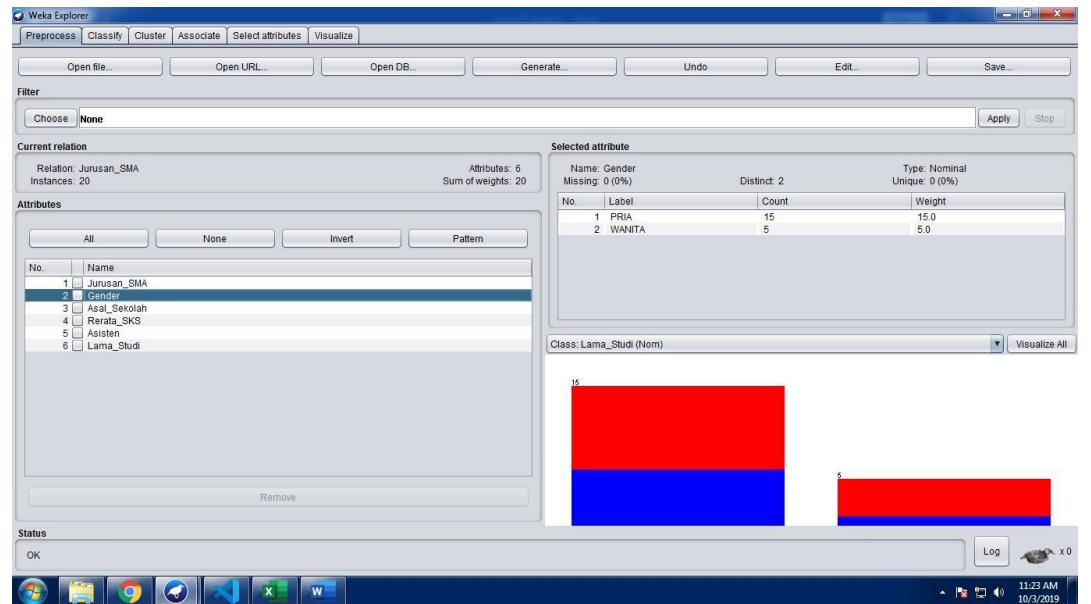
```
File Edit Selection View Go Debug Terminal Help
Welcome Jurusan_SMA.arff
C:\Users\LABSI-20\Documents\Jurusan_SMA.arff
@relation Jurusan_SMA
1 @attribute Jurusan_SMA {IPS,IPA,LAIN}
2
3 @attribute Gender {PRIA,WANITA}
4 @attribute Asal_Sekolah {SURAKARTA,LUAR}
5 @attribute Rerata_SKS REAL
6 @attribute Asisten {TIDAK,YA}
7 @attribute Lama_Studi {TERLAMBAT,TEPAT}
8 @attribute
9
10 @data
11 IPS,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
12 IPA,PRIA,SURAKARTA,19,YA,TERLAMBAT
13 LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
14 IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT
15 IPA,WANITA,SURAKARTA,17,TIDAK,TERPAT
16 IPA,WANITA,LUAR 18,YA,TERPAT
17 IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT
18 IPA,PRIA,SURAKARTA,19,TIDAK,TERPAT
19 IPS,PRIA,LUAR,18,TIDAK,TERLAMBAT
20 LAIN,WANITA,SURAKARTA,18,TIDAK,TERPAT
21 IPA,WANITA,SURAKARTA,19,TIDAK,TERPAT
22 IPS,PRIA,SURAKARTA,20,TIDAK,TERPAT
23 IPS,PRIA,SURAKARTA,19,TIDAK,TERPAT
24 IPA,PRIA,SURAKARTA,19,TIDAK,TERPAT
25 IPA,PRIA,LUAR,22,YA,TERPAT
26 LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT
27 IPS,PRIA,LUAR,20,TIDAK,TERPAT
28 LAIN,PRIA,LUAR,23,YA,TERPAT
29 IPA,PRIA,SURAKARTA,21,YA,TERPAT
30 IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
31
```

2. Gambar grafik dari WEKA

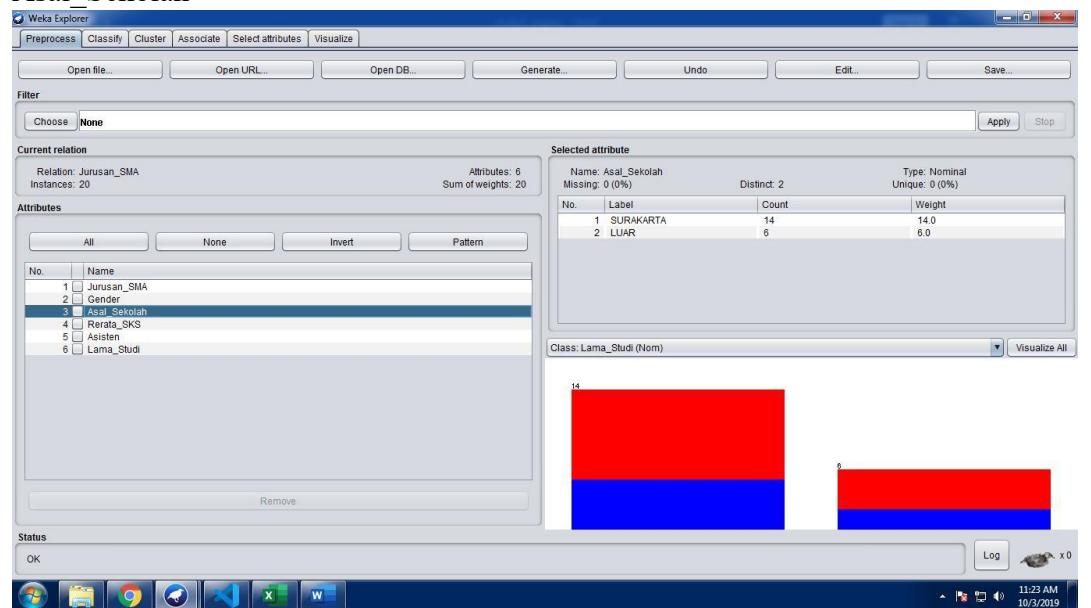
➤ Jurusan_SMA



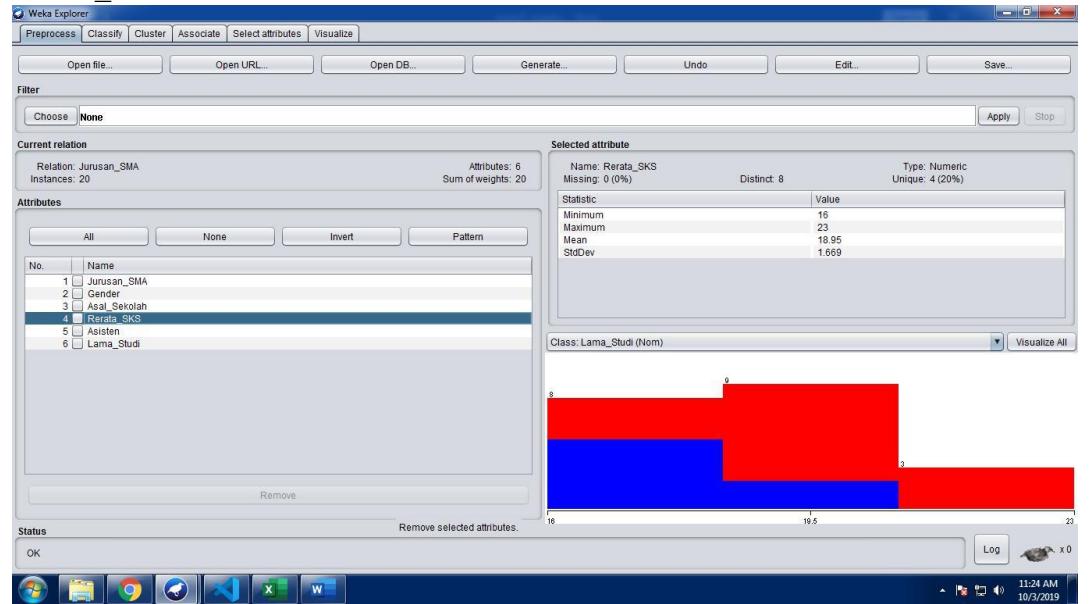
➤ Gender



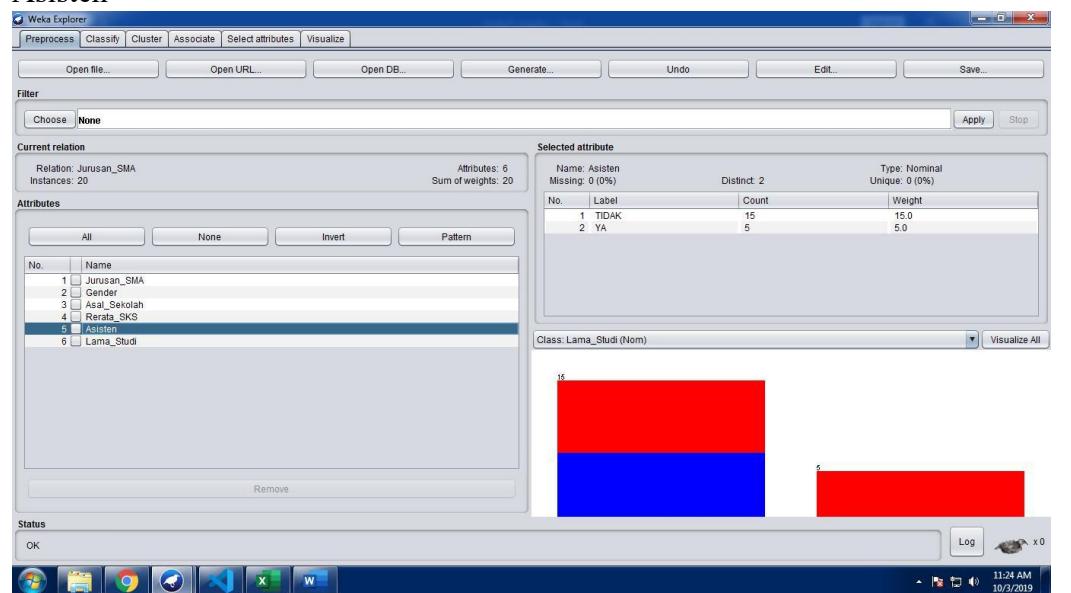
➤ Asal_Sekolah



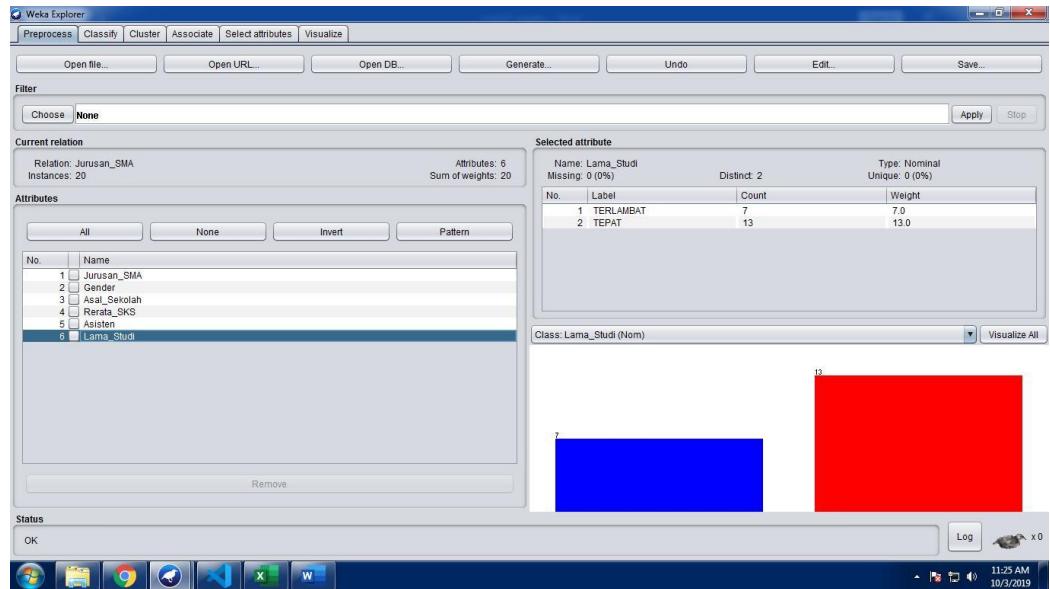
➤ Rerata_SKS



➤ Asisten



➤ Lama Studi



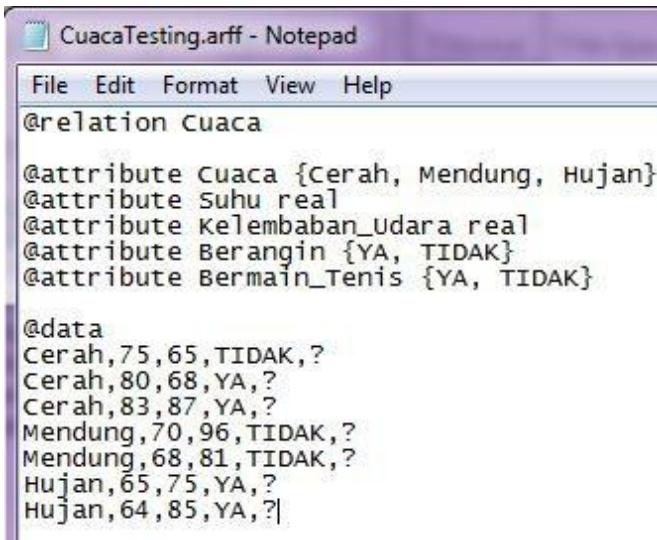
MODUL 8

Kegiatan

- Persiapkan file 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
```

- Buatlah seperti pada gambar dan save dengan nama CuacaTesting.arff:

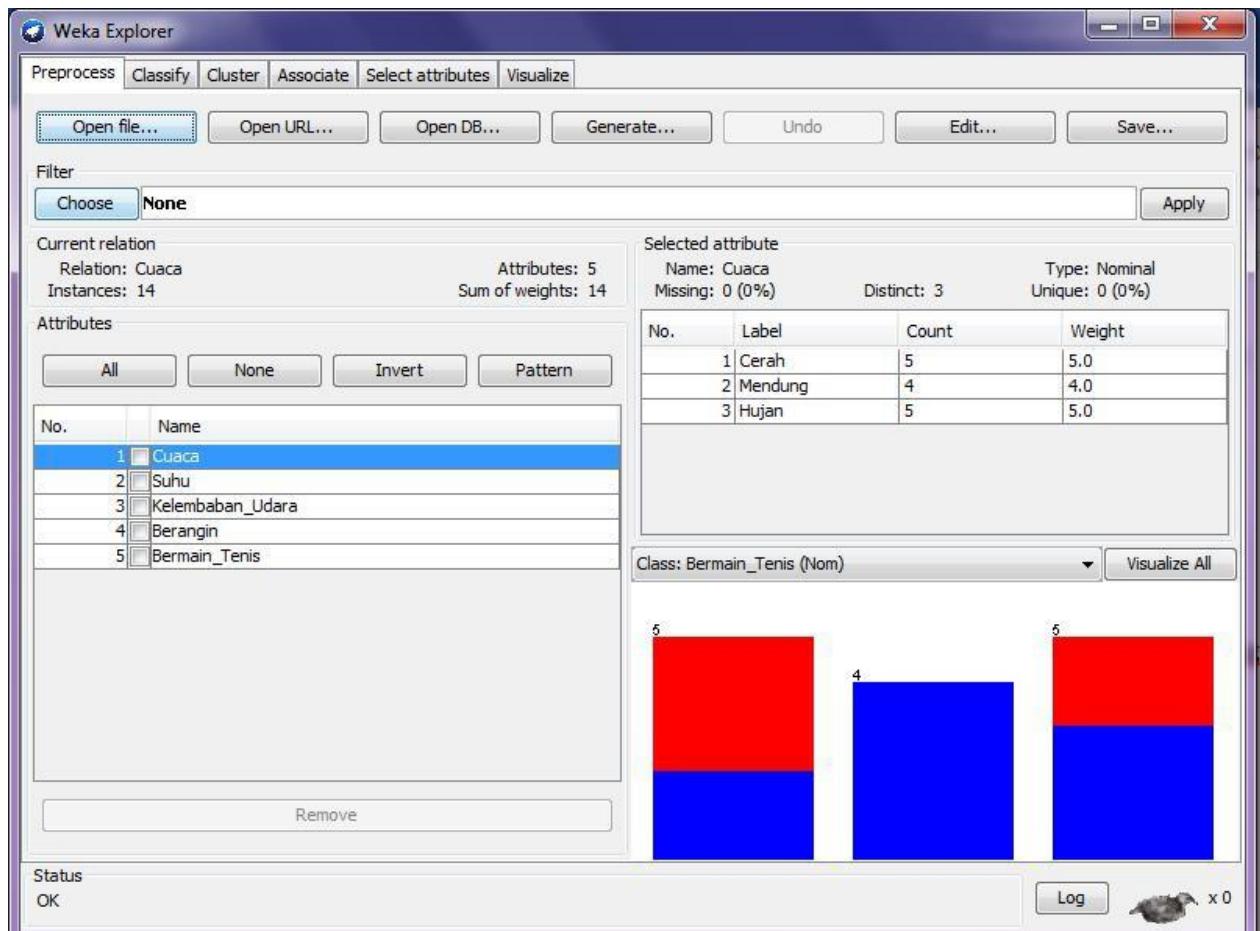


```
CuacaTesting.arff - Notepad
File Edit Format View Help
@relation Cuaca

@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_Udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

@data
Cerah,75,65,TIDAK,?
Cerah,80,68,YA,?
Cerah,83,87,YA,?
Mendung,70,96,TIDAK,?
Mendung,68,81,TIDAK,?
Hujan,65,75,YA,?
Hujan,64,85,YA,?|
```

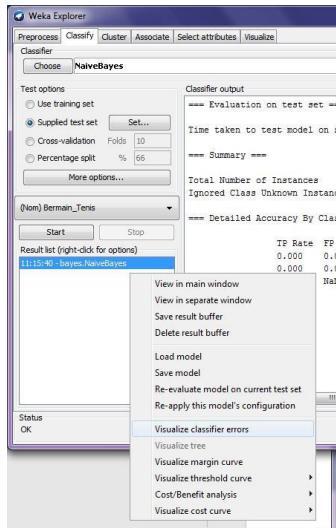
- Buka kembali file Cuaca.arff dengan menggunakan Weka Explorer. Buka Weka Explorer – Klik Open File – Pilih Cuaca.arff



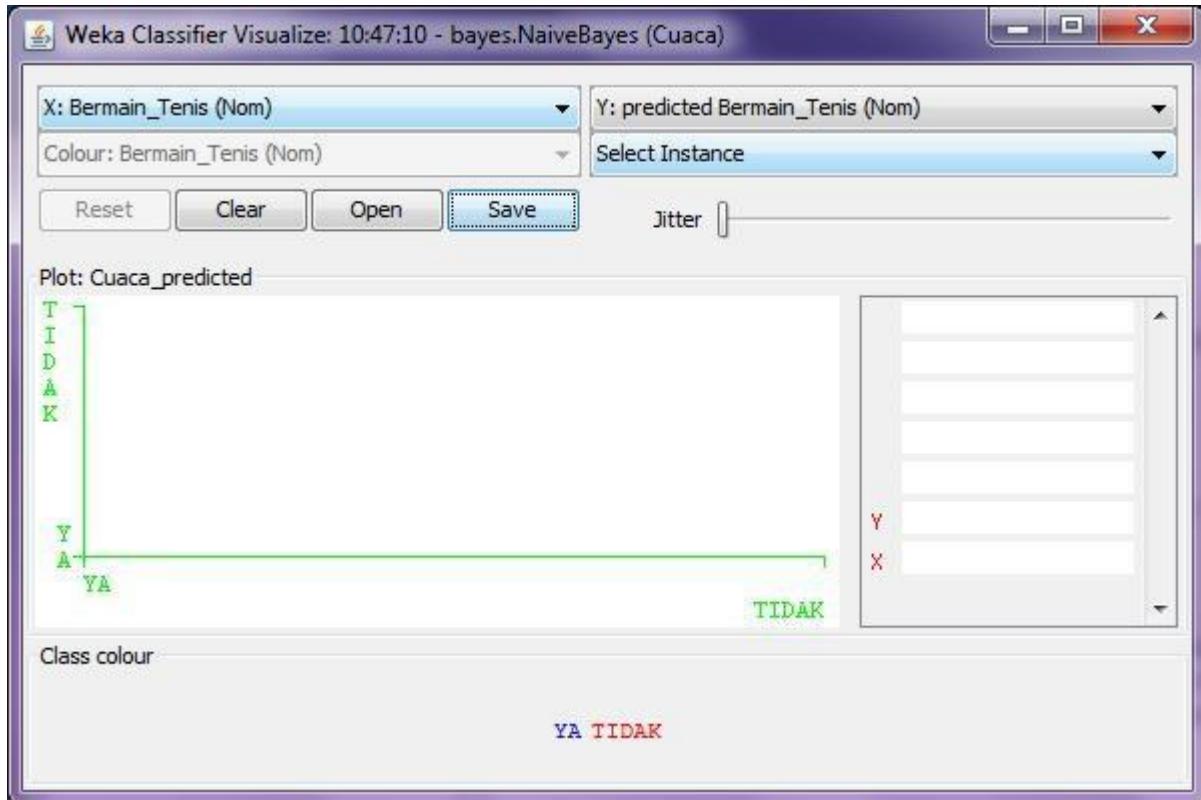
- Masih pada Weka Explorer, pilih tab **Classify**
- Sehingga akan muncul jendela Weka Explorer pada tab Classify, pada kotak **Classifier** klik tombol **Choose** untuk memilih metode/algoritma **Naïve Bayes**
- Klik tombol Set untuk menentukan file ARFF sebagai data uji
- Sehingga akan muncul jendela Test Instance, klik **Open file . . .**



8. Pilih file CuacaTesting.arff
9. Kemudian klik Close
10. Klik Start untuk memulai proses naïve bayes
11. Klik kanan pada hasil proses dalam kotak **result list**. Pilih **Visual classifier errors**



12. Klik Save. Simpan dengan nama file **HasilPrediksi.arff**



13. Tutup semua jendela termasuk Weka Explorer dan kembali ke Weka GUI Chooser. Pilih menu **Tools-ArffViewer**

14. Buka menu File-Open. Tunjukkan pada file HasilPrediksi.arff yang telah Anda simpan

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	

Implementasi Naïve Bayes dengan RapidMiner

- Persiapkan file **Tabel_Cuaca.xls** yang terdiri dari 2 sheet.
- Tabel data training** pada Sheet1

	A	B	C	D	E
1	Cuaca	Suhu	Kelembapan Udara	Berangin	Bermain_Tenis
2	Cerah	85	85	TIDAK	TIDAK
3	Cerah	80	90	YA	TIDAK
4	Mendung	83	86	TIDAK	YA
5	Hujan	70	96	TIDAK	YA
6	Hujan	68	80	TIDAK	YA
7	Hujan	65	70	YA	TIDAK
8	Mendung	64	65	YA	YA
9	Cerah	72	95	TIDAK	TIDAK
10	Cerah	69	70	TIDAK	YA
11	Hujan	75	80	TIDAK	YA
12	Cerah	75	70	YA	YA
13	Mendung	72	90	YA	YA
14	Mendung	81	75	TIDAK	YA
15	Hujan	71	91	YA	TIDAK

3. Tabel data uji pada Sheet2 tanpa ada variabel **Bermain_Tenis**

	A	B	C	D
1	Cuaca	Suhu	Kelembapan_udara	Berangin
2	Cerah		75	65 TIDAK
3	Cerah		80	68 YA
4	Cerah		83	87 YA
5	Mendung		70	96 TIDAK
6	Mendung		68	81 TIDAK
7	Hujan		65	75 YA
8	Hujan		64	85 YA

4. Buka aplikasi Rapid Miner
5. Klik **Import Data**. Arahkan direktori tempat penyimpanan file pada langkah **Select data location**, kemudian pilih file yang akan digunakan dan klik **Next**.
6. Pastikan sel Excel sesuai di langkah **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	F
1 Cuaca	Suhu	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	

Ac ← Previous → Next ✖ Cancel

7. Pada langkah **Format your columns** ubah kolom **Bermain_Tenis** dengan tipe data **binomial** karena hanya ada dua keputusan (YA dan TIDAK)

Format your columns.

Replace errors with missing values ⓘ

Cuaca	Suhu	Kelembaban_Udara	Berangin	Bermain_Tenis
polynomial	integer	integer	polynomial	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

Change Type ▶ polynomial
 binomial
 real
 integer
 date_time
 date
 time

8. Ubah pula sebagai **label** pada Change Role

The screenshot shows the KNIME interface with a data table titled "Bermain_Tenis binomial". A context menu is open over the first column, with "Change Role" selected. The menu also includes "Re-Opens a dialog to change the role." and "Exclude column". To the right, the modified data table is shown with the first column labeled "label" and the value "TIDAK" highlighted in green.

9. Simpan dengan nama **DataCuaca_Training** dilanjutkan klik tombol **Finish**

10. Hasil import file **Tabel_Cuaca.xls** pada Sheet1 akan di tampilkan.

The screenshot shows the KNIME interface with the title "ExampleSet //Local Repository/DataCuaca_Training". Below it, there are buttons for "Open in Turbo Prep" and "Auto Model". A filter bar indicates "Filter (14 / 14 examples): all". The main area displays a data table with 14 rows and 6 columns:

Row No.	Bermain_Tenis	Cuaca	Suhu	Kelembaban_Udara	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

11. Kembali ke jendela Design Perspective dengan shortcut tombol **F8**

12. Lakukan hal yang sama untuk data testing yang diambil dari **Tabel_Cuaca.xls** pada Sheet2(Testing) dengan mengulang dari langkah 5

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

13. Simpan dengan nama **DataCuaca_Testing**.

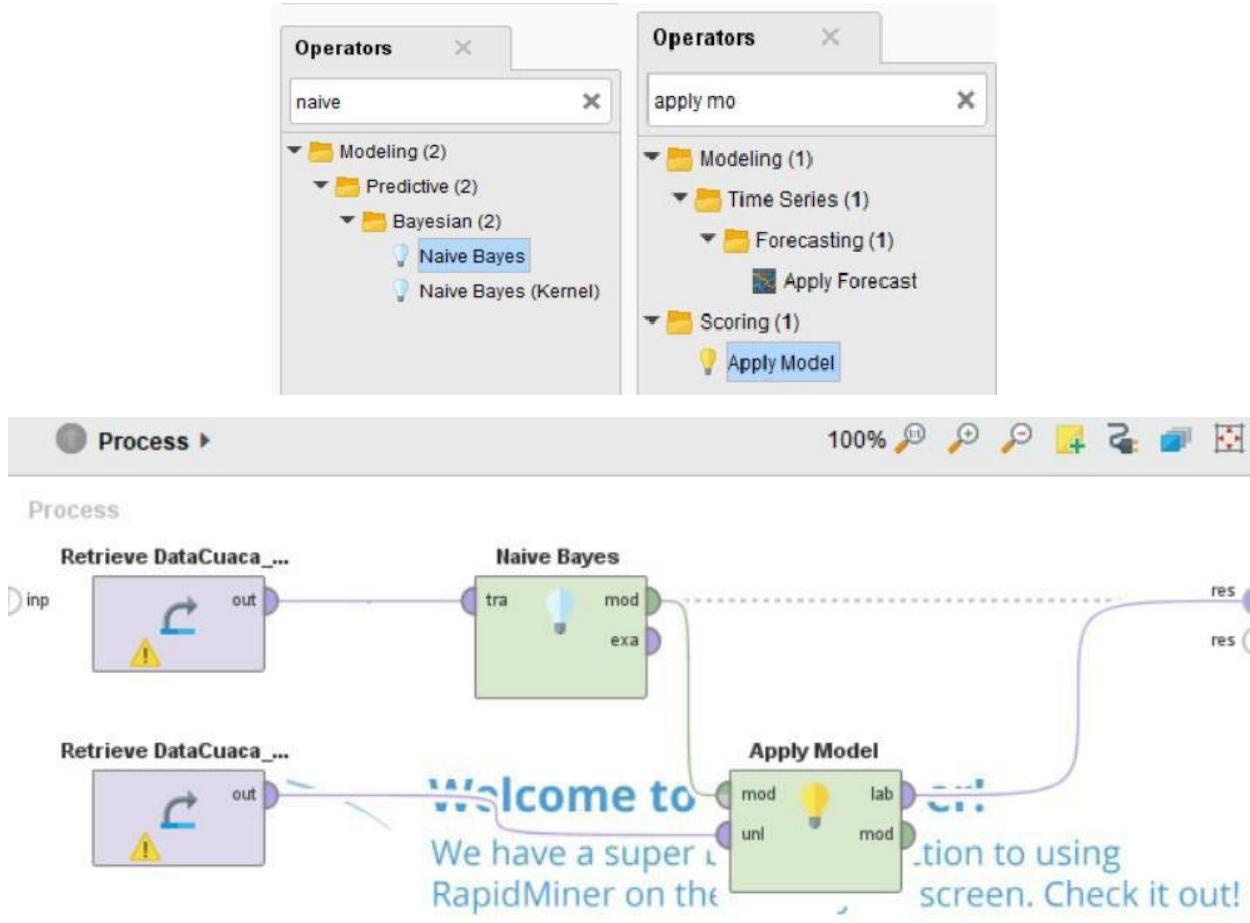
ExampleSet //Local Repository/DataCuaca_Testing

Open in [Turbo Prep](#) [Auto Model](#) Filter (7 / 7 examples):

Row No.	Cuaca	Suhu	Kelembaban_Udara	Berangin
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

14. Drag **DataCuaca_Training** dan **DataCuaca_Testing** kedalam jendela Process View

15. Masukkan juga operator **Naive Bayes** dan **Apply Model** ke dalam Proses View.
 Hubungkan koneksi masing-masing data terhadap operator seperti gambar:



16. Jalankan proses naive bayes dengan menekan tombol **Run** (atau menekan tombol F11).
 17. Perhatikan hasil proses klasifikasi naive bayes. Pada tab **Data**, dapat dilihat hasil prediksi serta tingkat confidence nilai kelas pada masing-masing data.

Row No.	prediction(B...)	confidence(...)	confidence(...)	Cuaca	Suhu	Kelembaban...	Berangin
1	YA	0.154	0.846	Cerah	75	65	TIDAK
2	YA	0.498	0.502	Cerah	80	68	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	65	75	YA
7	TIDAK	0.568	0.432	Hujan	64	85	YA

Pada tab **Statistics**, dapat dilihat bahwa distribusi nilai kelas pada variabel Y (Bermain_Tenis) rerata nilai confidence sebesar 0,353 untuk nilai TIDAK, dan 0,647 untuk nilai YA

	Name	Type	Missing	Statistics	Filter (7 / 7 attributes): <input type="text" value="Search for Attributes"/>
Data	prediction(Bermain_Tenis)	Binominal	0	Least TIDAK (2) Most YA (5) Values YA (5), TIDAK (2)	
Statistics	confidence(TIDAK)	Real	0	Min 0.007 Max 0.856 Average 0.353	
Visualizations	confidence(YA)	Real	0	Min 0.144 Max 0.993 Average 0.647	
Annotations	Cuaca	Polynominal	0	Least Mendung (2) Most Cerah (3) Values Cerah (3), Hujan (2), ...[1]	
	Suhu	Integer	0	Min 64 Max 83 Average 72.143	
	Kelembaban_Udara	Integer	0	Min 65 Max 96 Average 79.571	
	Berangin	Polynominal	0	Least TIDAK (3) Most YA (4) Values YA (4), TIDAK (3)	

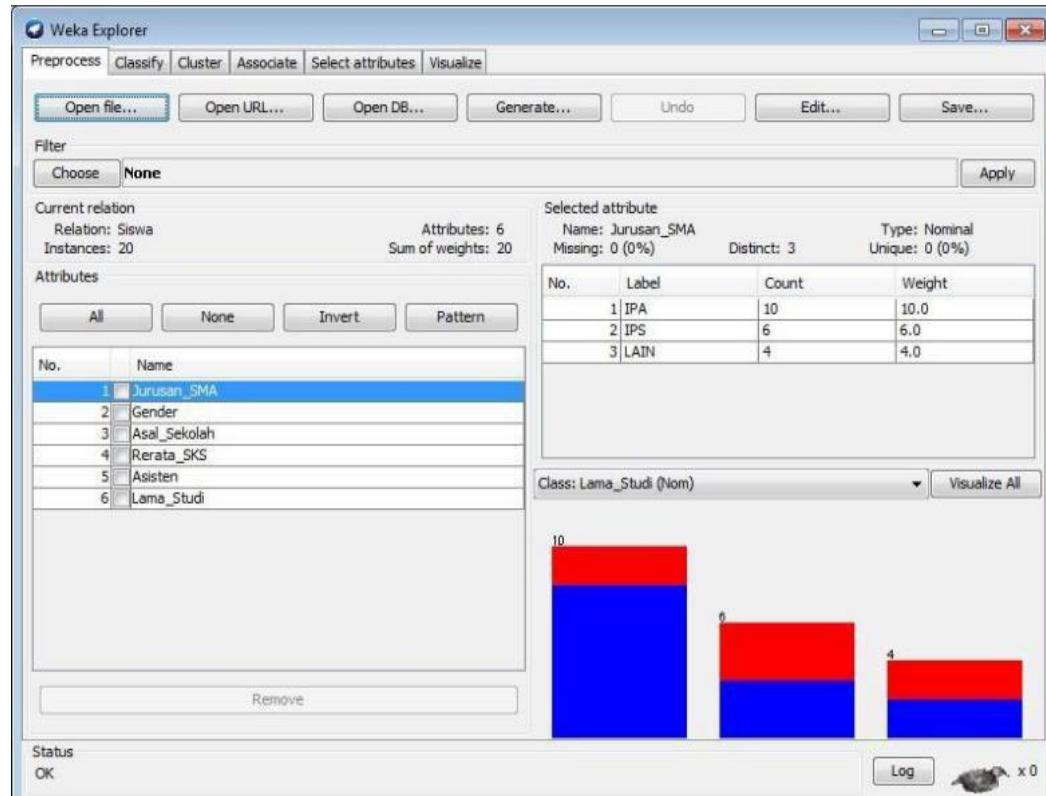
Tugas

1. Berdasarkan tabel berikut, buatlah file dalam format Excel (.xls) dan format ARFF (.arff). Data ini akan digunakan sebagai **Data Testing**.

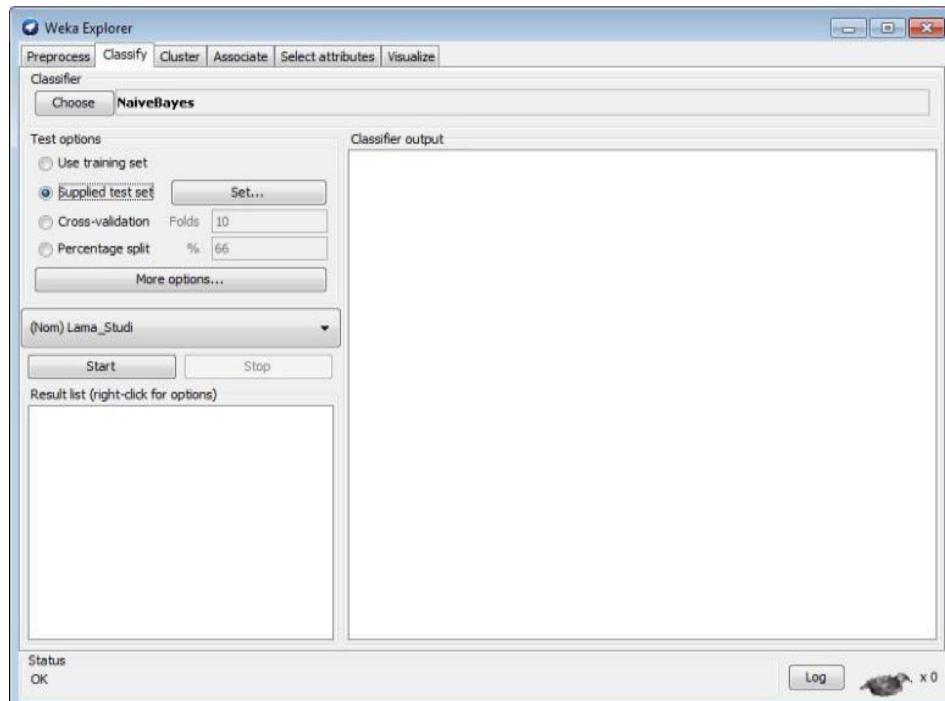
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	WANITA	SURAKARTA	18	TIDAK
3	IPA	PRIA	SURAKARTA	19	YA
4	LAIN	PRIA	SURAKARTA	19	TIDAK
5	IPS	PRIA	LUAR	17	TIDAK
6	LAIN	WANITA	SURAKARTA	17	TIDAK
7	IPA	WANITA	LUAR	18	YA
8	IPA	PRIA	SURAKARTA	18	TIDAK
9	IPA	PRIA	SURAKARTA	19	TIDAK
10	IPS	PRIA	LUAR	18	TIDAK
11	LAIN	WANITA	SURAKARTA	18	TIDAK

2. Gunakan file ARFF yang dikerjakan pada Tugas nomor 1 dalam Modul 7 sebagai data training. Lakukan prediksi terhadap data testing (ARFF) di atas menggunakan WEKA!

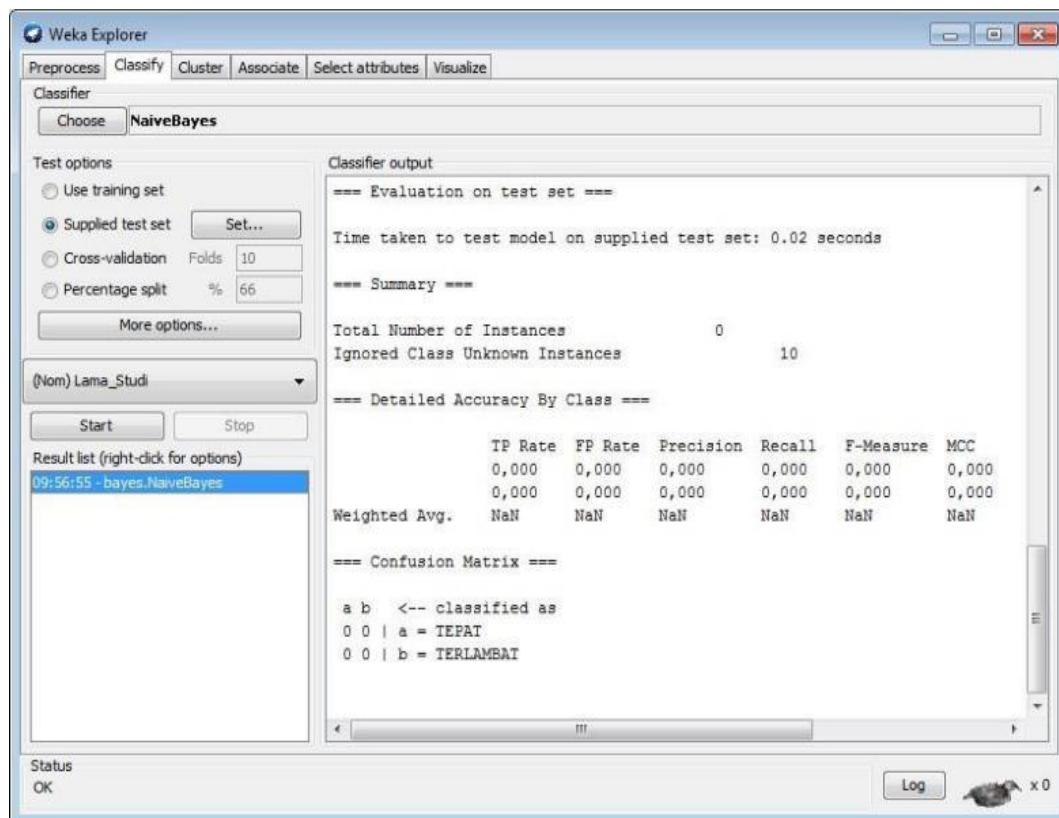
- Membuat File Training



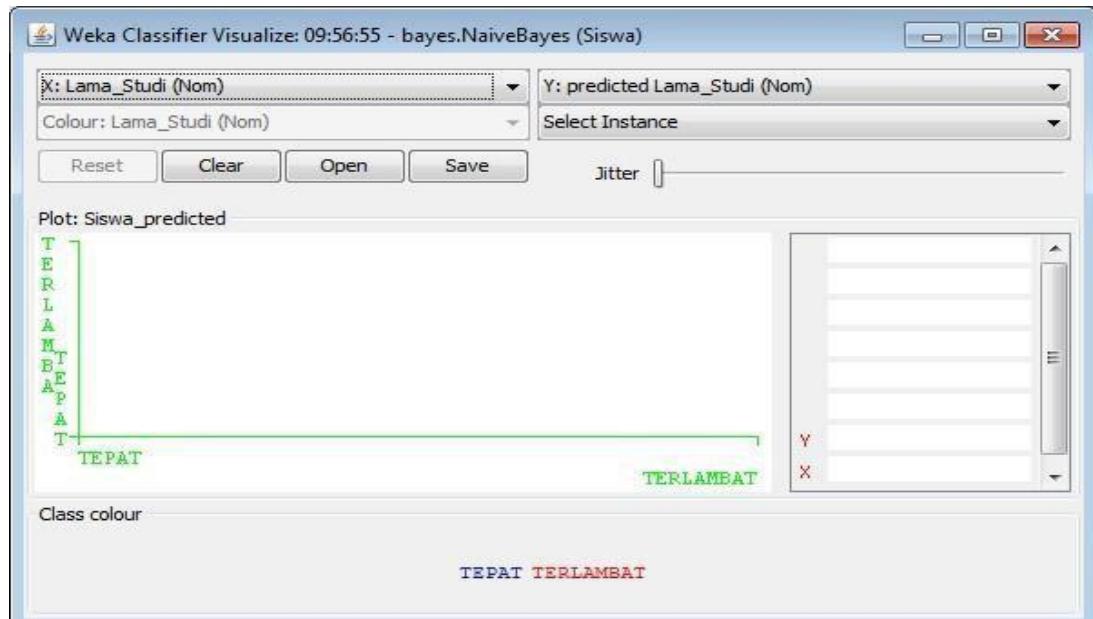
- Memilih Metode Naive Bayes



- Membuka file testing lalu klik kanan pilih visualize classifier errors



- Menyimpan file hasil



- Hasil Prediksi

The screenshot shows the ARFF-Viewer application displaying a table of student data. The title bar says 'ARFF-Viewer - D:\dwdm\modul 8\tugas\HasiTugas.arff'. The menu bar includes 'File', 'Edit', and 'View'. The table has a header row with columns: No., 1: Jurusan_SMA Nominal, 2: Gender Nominal, 3: Asal_Sekolah Nominal, 4: Rerata_SKS Numeric, 5: Asisten Nominal, 6: prediction margin Numeric, 7: predicted Lama_Studi Nominal, and 8: Lama_Studi Nominal. Below the header, there are 10 data rows. The 'predicted Lama_Studi Nominal' column contains values like 'TERLAMBAT' and 'TEPAT', which correspond to the green and red bars shown in the Weka plot above.

No.	1: Jurusan_SMA 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	

3. Gunakan file Excel yang dikerjakan pada Tugas nomor 1 dalam Modul 6 sebagai data training. Lakukan prediksi terharap data testing (Excel) di atas menggunakan RapidMiner!

- Data Training

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet1 ▾ Cell range: A:F Select All Define header row: 1

	A	B	C	D	E	F
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
2	IPS	WANITA	SURAKARTA	18.000	TIDAK	TERLAMBAT
3	IPA	PRIA	SURAKARTA	19.000	YA	TEPAT
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK	TERLAMBAT
5	IPA	PRIA	LUAR	17.000	TIDAK	TERLAMBAT
6	IPA	WANITA	SURAKARTA	17.000	TIDAK	TEPAT
7	IPA	WANITA	LUAR	18.000	YA	TEPAT
8	IPA	PRIA	SURAKARTA	18.000	TIDAK	TERLAMBAT
9	IPA	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
10	IPS	PRIA	LUAR	18.000	TIDAK	TERLAMBAT
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK	TEPAT
12	IPA	WANITA	SURAKARTA	19.000	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	20.000	TIDAK	TEPAT
14	IPS	PRIA	SURAKARTA	19.000	TIDAK	TEPAT
15	IPA	PRIA	SURAKARTA	18.000	TIDAK	TEPAT

← Previous → Next X Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ?

	Jurusan_SMA * <small>polynomial</small>	Gender * <small>polynomial</small>	Asal_Sekolah * <small>polynomial</small>	Rerata_SKS * <small>integer</small>	Asisten <small>binomial</small>	Lama_Studi * <small>binomial 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	SIRAKARTA	19	TIDAK	TEPAT

✓ no problems.

← Previous Next → Cancel X

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Find data, operators, etc. All Studio

ExampleSet ((Local Repository/DataCuaca_Testing)) ExampleSet ((Local Repository/DataCuaca_Training)) ExampleSet (Apply Model)

Result History ExampleSet ((Local Repository/DataSekolah_Training)) ExampleSet (Apply Model)

Data Statistics Visualizations Annotations

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)

Repository

- Import Data
- Samples
- DB (empty)
- Local Repository (empty)
 - Connections (empty)
 - data (empty)
 - processes (empty)
 - DataCuaca_Testing (data - v1, 19/01/19 10:19:16)
 - DataCuaca_Training (data - v1, 19/01/19 10:19:16)
 - DataSekolah_Training (data - v1, 19/01/19 10:19:16)
 - HasilCuaca (data - v1, 19/01/19 10:00:00)

- Data Testing

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet2 ▾ Cell range: A:E Select All Define header row: 1

	A	B	C	D	E
1	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
2	LAIN	WANITA	SURAKARTA	18.000	TIDAK
3	IPA	PRIA	SURAKARTA	19.000	YA
4	LAIN	PRIA	SURAKARTA	19.000	TIDAK
5	IPS	PRIA	LUAR	17.000	TIDAK
6	LAIN	WANITA	SURAKARTA	17.000	TIDAK
7	IPA	WANITA	LUAR	18.000	YA
8	IPA	PRIA	SURAKARTA	18.000	TIDAK
9	IPA	PRIA	SURAKARTA	19.000	TIDAK
10	IPS	PRIA	LUAR	18.000	TIDAK
11	LAIN	WANITA	SURAKARTA	18.000	TIDAK

← Previous → Next ✖ Cancel

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Jurusan_SMA polynomial	Gender polynomial	Asal_Sekolah polynomial	Rerata_SKS integer	Asisten polynomial
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

✓ no problems.

← Previous → Next ✖ Cancel

RapidMiner Studio Educational 9.3.001 @ asus-PC

Views: Design Results Turbo Prep Auto Model

Find data, operators...etc All Studio

Repository

Import Data

Local Repository (asus)

- Samples
- DB (Legacy)
- processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9:25 AM)
 - DataCuaca_Training (asus - v1, 10/5/19 9:25 AM)
 - DataSekolah_Testing (asus - v1, 10/5/19 9:25 AM)
 - DataSekolah_Training (asus - v1, 10/5/19 9:25 AM)
 - HasilCuaca (asus - v1, 10/5/19 10:00 AM -)

Data

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

ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)

ExampleSet (/Local Repository/DataSekolah_Testing)

Row History

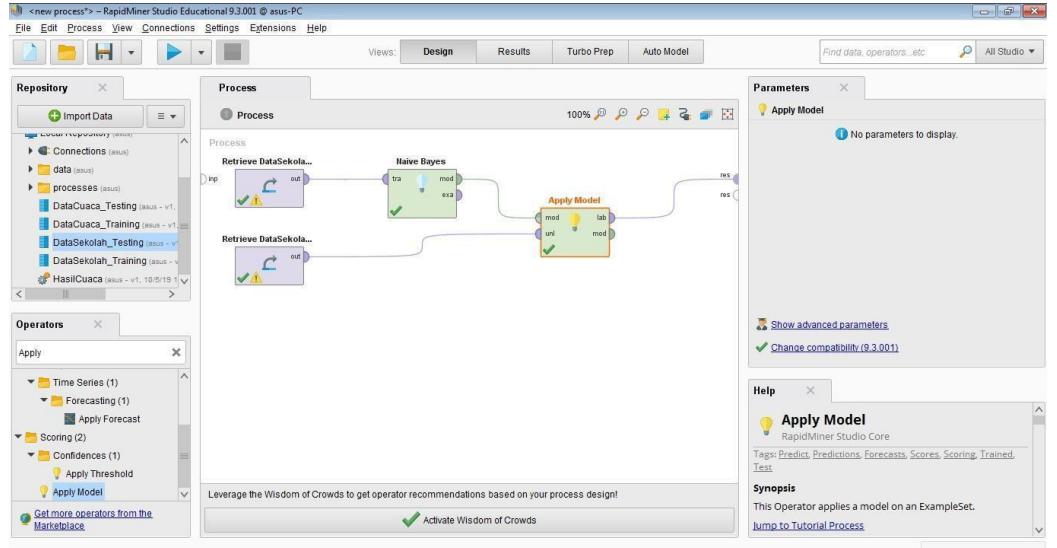
Open in Turbo Prep Auto Model

Filter (10 / 10 examples): all

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
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, 5 regular attributes)

- **Implementasi Naive Bayes**



- Hasil

Screenshot 1: RapidMiner Studio Results View (Top)

This screenshot shows the RapidMiner Studio interface with the 'Results' tab selected. The main area displays a table titled 'ExampleSet (Apply Model)' containing 10 examples. The columns include Row No., prediction(Lama_Studi), confidence_(TERLAMBAT), confidence_(TEPAT), Jurusan_SMA, Gender, Asal_Sekolah, Rerata_SKS, and Asisten.

Row No.	prediction(Lama_Studi)	confidence_(TERLAMBAT)	confidence_(TEPAT)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK
2	TEPAT	0.005	0.995	IPA	PRIA	SURAKARTA	19	YA
3	TERLAMBAT	0.650	0.350	LAIN	PRIA	SURAKARTA	19	TIDAK
4	TERLAMBAT	0.888	0.132	IPS	PRIA	LUAR	17	TIDAK
5	TERLAMBAT	0.738	0.262	LAIN	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	0.811	0.189	IPS	PRIA	LUAR	18	TIDAK
10	TERLAMBAT	0.648	0.352	LAIN	WANITA	SURAKARTA	18	TIDAK

The repository sidebar on the right shows the local repository with datasets like DataCuaca_Testing, DataCuaca_Training, DataSekolah_Testing, and DataSekolah_Training.

Screenshot 2: RapidMiner Studio Results View (Bottom)

This screenshot shows the same RapidMiner Studio interface, but the results view has been expanded to show detailed statistics for each attribute. The left sidebar shows 'Annotations' selected.

Name	Type	Missing	Statistics	Filter (0 / 8 attributes)	Search for Attributes
prediction(Lama_Studi)	Binomial	0	Least: TEPAT (3) Most: TERLAMBAT (7) Values: TERLAMBAT (7), TEPAT (3)		
confidence_(TERLAMBAT)	Real	0	Min: 0.005 Max: 0.868 Average: 0.524		
confidence_(TEPAT)	Real	0	Min: 0.132 Max: 0.995 Average: 0.476		
Jurusan_SMA	Polynominal	0	Least: IPS (2) Most: IPA (4) Values: IPA (4), LAIN (4), ... [more]		
Gender	Polynominal	0	Least: WANITA (4) Most: PRIA (6) Values: PRIA (6), WANITA (4)		
Asal_Sekolah	Polynominal	0	Least: LUAR (3) Most: SURAKARTA (7) Values: SURAKARTA (7), LUAR (3)		
Rerata_SKS	Integer	0	Min: 17 Max: 19 Average: 18.100		
Asisten	Polynominal	0	Least: YA (2) Most: TIDAK (8) Values: TIDAK (8), YA (2)		

The bottom status bar indicates 'Showing attributes 1 - 8' and 'Examples: 10 Special Attributes: 3 Regular Attributes: 5'.

4. Dari hasil percobaan Tugas nomor 3 di atas, berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT? Berapakah nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT?

Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.005	Max 0.868	Average 0.524
Confidence_TEPAT confidence(TEPAT)	Real	0	Min 0.132	Max 0.995	Average 0.476

Jadi nilai rerata confidence untuk atribut Lama_Studi dengan nilai TEPAT adalah **0,524** dan nilai rerata confidence untuk atribut Lama_Studi dengan nilai TERLAMBAT adalah **0,476**

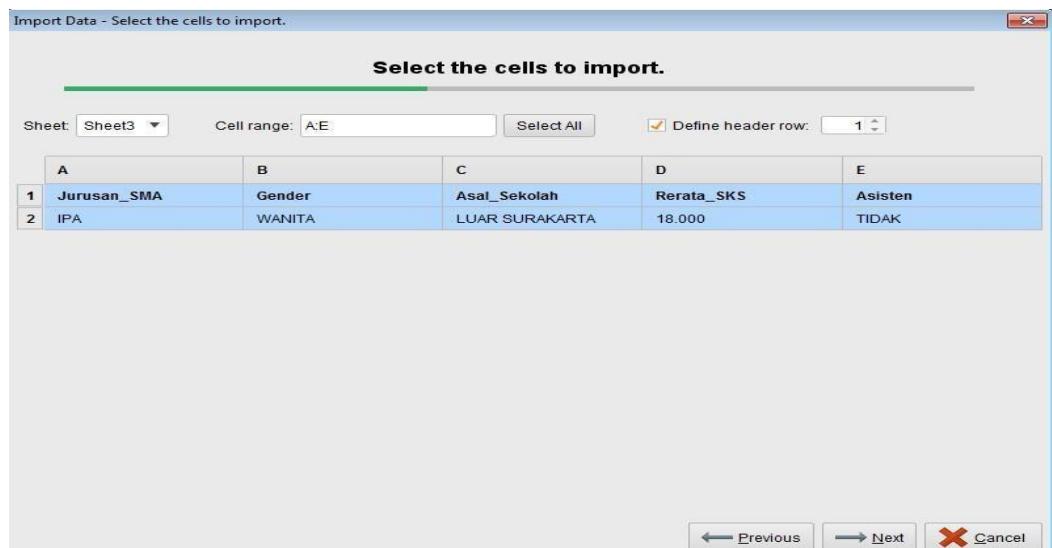
5. Dari hasil percobaan Tugas nomor 3 di atas, berapa orang yang akan lulus TEPAT, dan berapa orang yang akan lulus TERLAMBAT?

Prediction prediction(Lama_Studi)	Binomial	0	Least TEPAT (3)	Most TERLAMBAT (7)	Values TERLAMBAT (7), TEPAT (3)
--------------------------------------	----------	---	-----------------	--------------------	---------------------------------

Jadi yang lulus dengan tepat sebanyak **3** orang sedangkan yang lulus terlambat sebanyak **7** orang

6. Prediksi ketepatan lama studi si Dewi, jika Dewi adalah seorang WANITA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari LUAR SURAKARTA, mengambil SKS dengan rata-rata sebanyak 18 SKS tiap semester, dan tidak pernah menjadi Asisten selama kuliah.

- Mengimport Data Dewi



Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Jurusan_SMA polynominal	Gender polynominal	Asal_Sekolah polynominal	Rerata_SKS integer	Asisten polynominal
1	IPA	WANITA	LUAR SURAKARTA	18	TIDAK

PREVI... no problems.

Previous Next Cancel

Result History

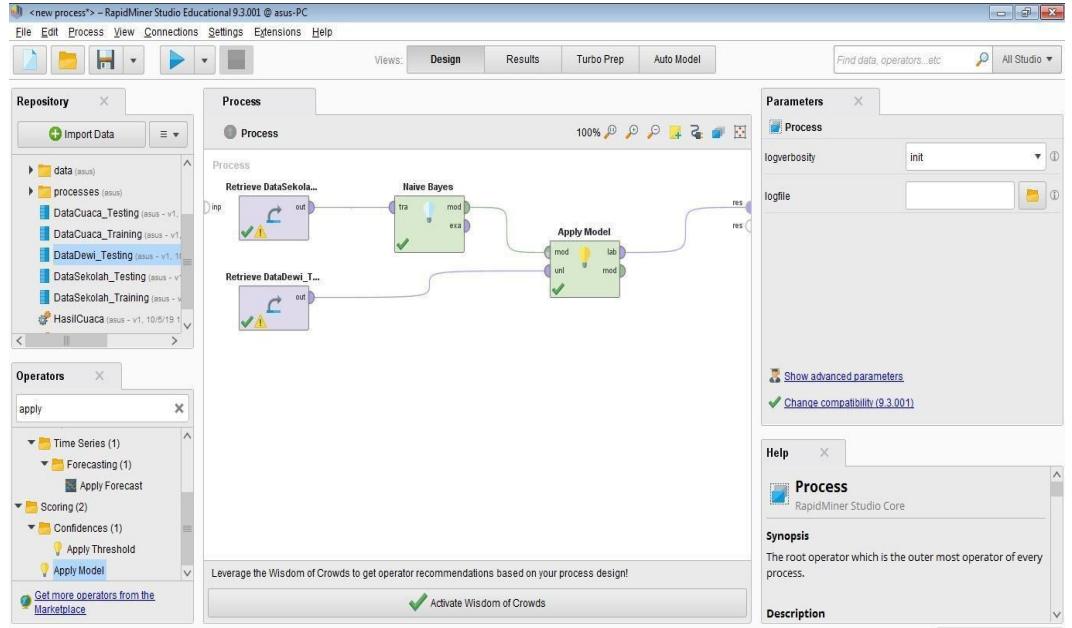
ExampleSet (/Local Repository/Data)

Open In Turbo Prep Auto Model

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	IPA	WANITA	LUAR SURA...	18	TIDAK

Data Statistics Visualizations Annotations

- Mengimplementasikan Naive Bayes



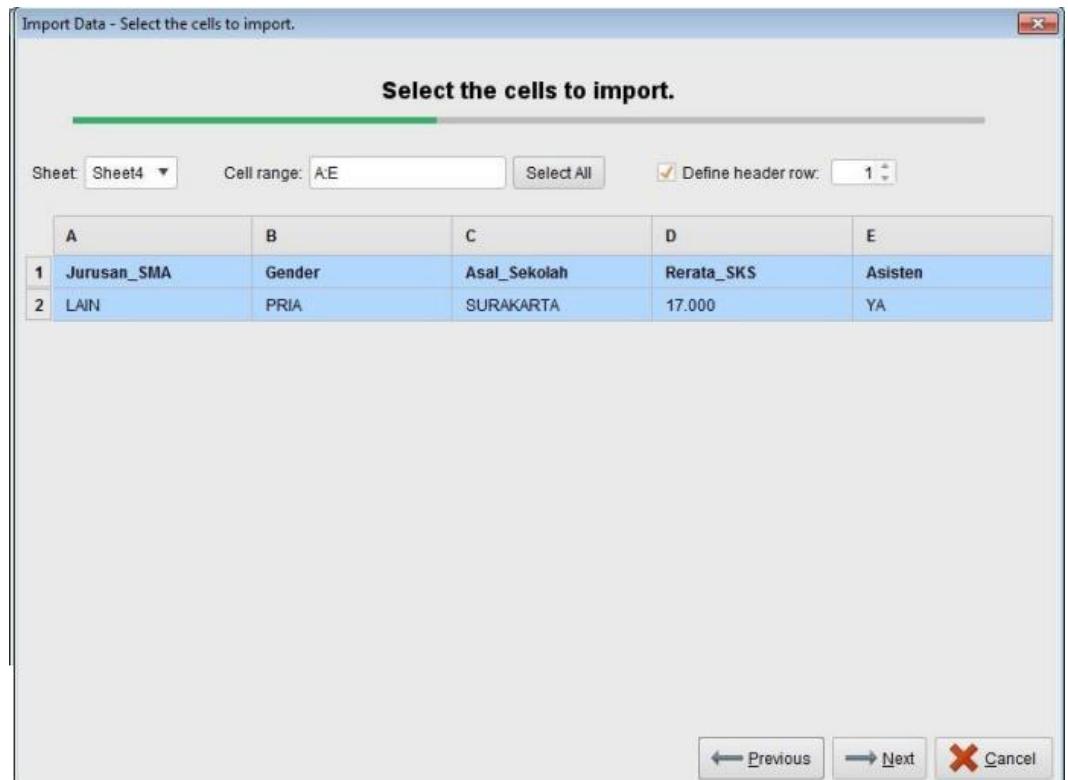
- Hasil

Row No.	prediction(L...)	confidence(...)	confidence(...)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.457	0.543	IPA	WANITA	LUAR SURAKARTA	18	TIDAK

Berdasarkan hasil di atas maka prediksi untuk Dewi dia akan lulus dengan tepat

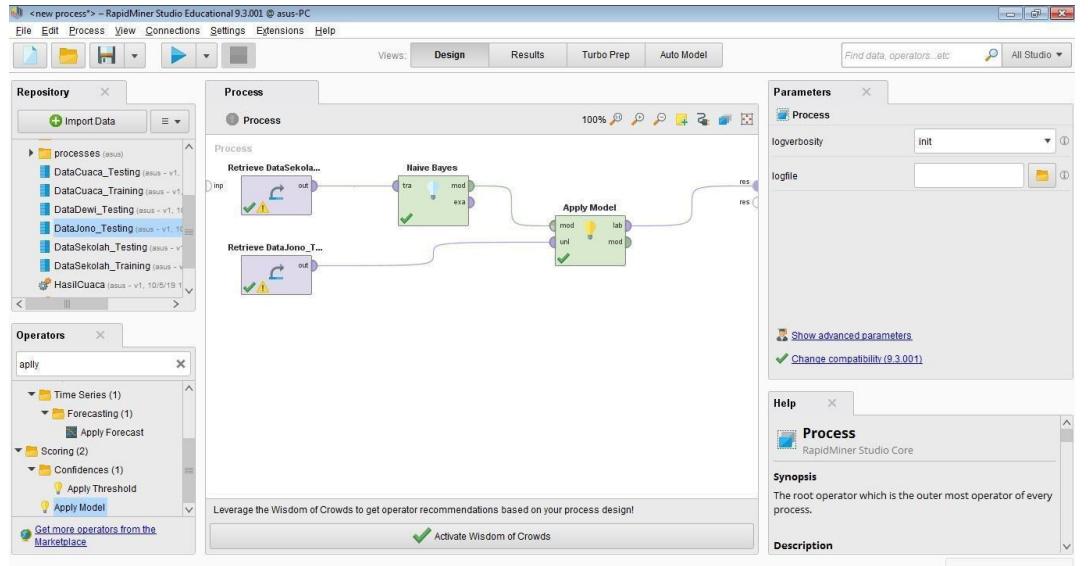
7. Prekdisikan ketepatan lama studi si Jono, jika Jono adalah seorang PRIA yang berasal dari jurusan IPA pada saat SMA, asal sekolah dari SURAKARTA, mengambil SKS dengan rata-rata sebanyak 17 SKS tiap semester, dan pernah menjadi Asisten selama kuliah.

- Mengimport Data Jono



Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	LAIN	PRIA	SURAKARTA	17	YA

- Mengimplementasikan Naïve Bayes



- Hasil

The screenshot shows the RapidMiner Studio interface with the "Results" view open. It displays the output of the "ExampleSet (Apply Model)" operator. The "Data" tab shows a single row with the following values:

Row No.	prediction(Lama_Studi)	confidence(TEPAT)	confidence(TERLAMBAT)	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

The "Statistics" tab shows the distribution of attributes:

Name	Type	Missing	Statistics
prediction(Lama_Studi)	Binomial	0	Least: TERLAMBAT (0), Most: TEPAT (1), Values: TEPAT (1), TERLAMBAT (0)
confidence(TERLAMBAT)	Real	0	Min: 0.076, Max: 0.076, Average: 0.076
confidence(TEPAT)	Real	0	Min: 0.924, Max: 0.924, Average: 0.924
Jurusan_SMA	Polynomial	0	Least: LAIN (1), Most: LAIN (1), Values: LAIN (1)
Gender	Polynomial	0	Least: PRIA (1), Most: PRIA (1), Values: PRIA (1)
Asal_Sekolah	Polynomial	0	Least: SURAKARTA (1), Most: SURAKARTA (1), Values: SURAKARTA (1)

The "Annotations" tab shows the predicted values for the attributes:

Name	Type	Missing	Statistics
Rerata_SKS	integer	0	Min: 17, Max: 17, Average: 17
Asisten	Polynomial	0	Least: YA (1), Most: YA (1), Values: YA (1)

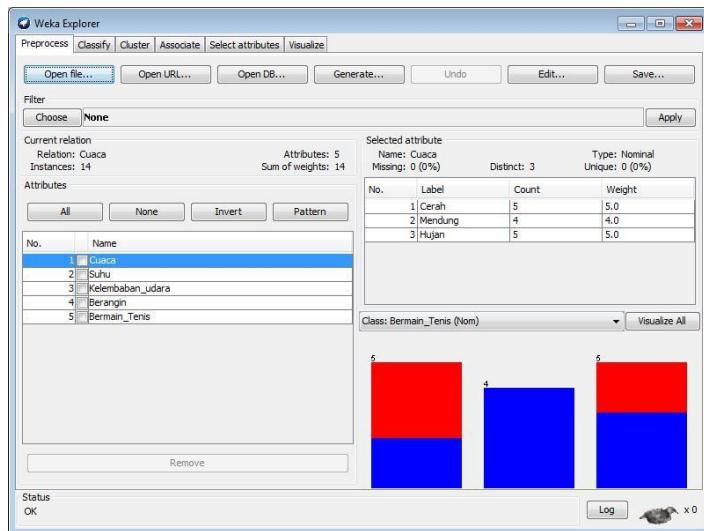
Berdasarkan hasil di atas maka prediksi untuk Jono dia akan lulus dengan tepat

MODUL 9

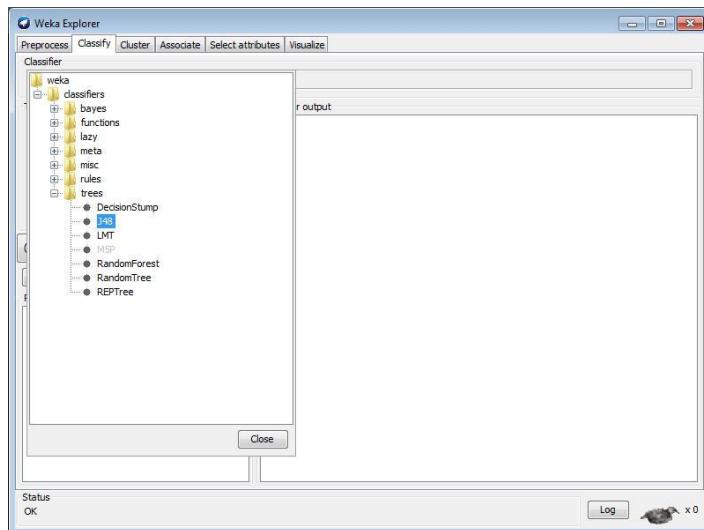
Kegiatan

Menggunakan WEKA

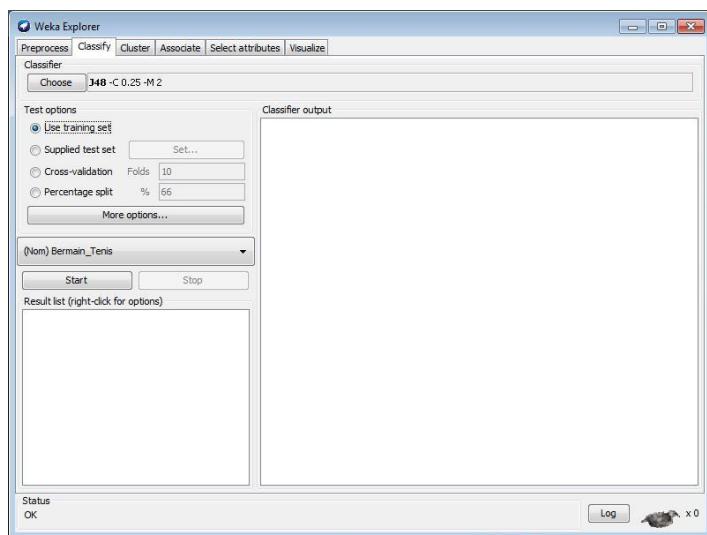
1. Membuka file Cuaca.arff dengan WEKA



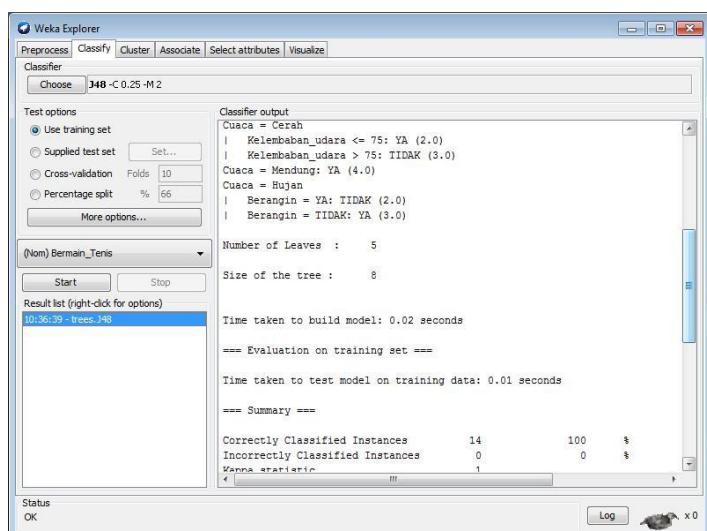
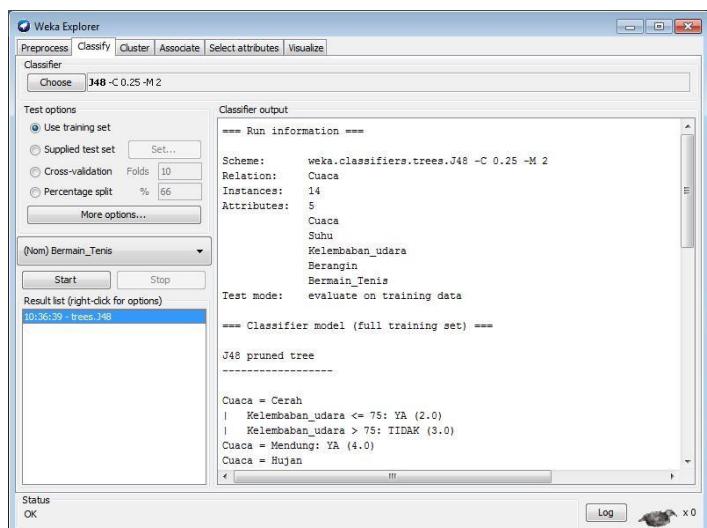
2. Membuka tab Classify dan tekan tombol choose-trees-j48

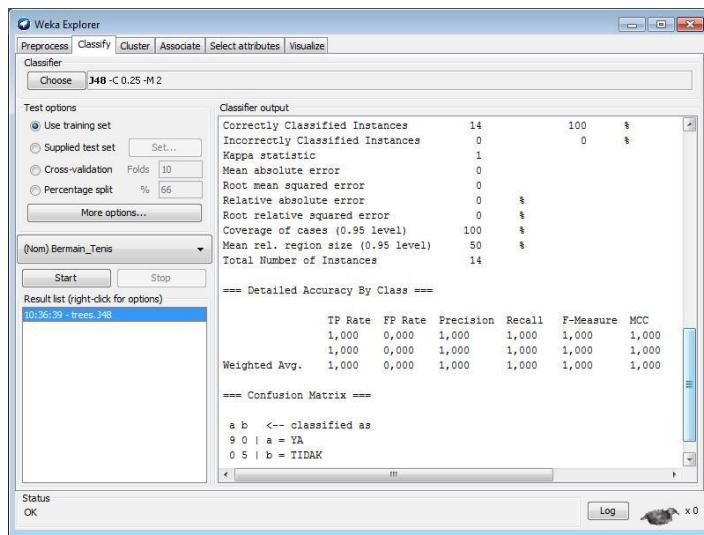


3. Pada pilihan Test Options, pilih Use training set. Kemudian klik start.

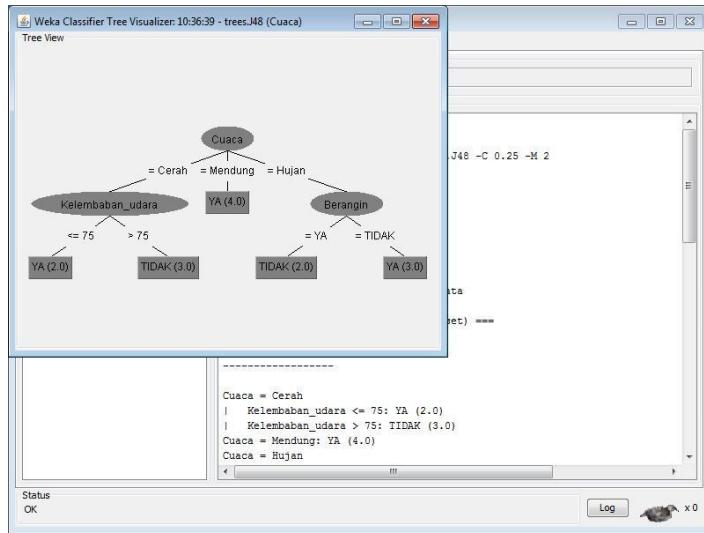


4. Hasil perhitungan klasifikasi menggunakan algoritma j48.





5. Hasil klasifikasi pohon keputusan.



Menggunakan Rapidminer

1. DataCuaca_Training

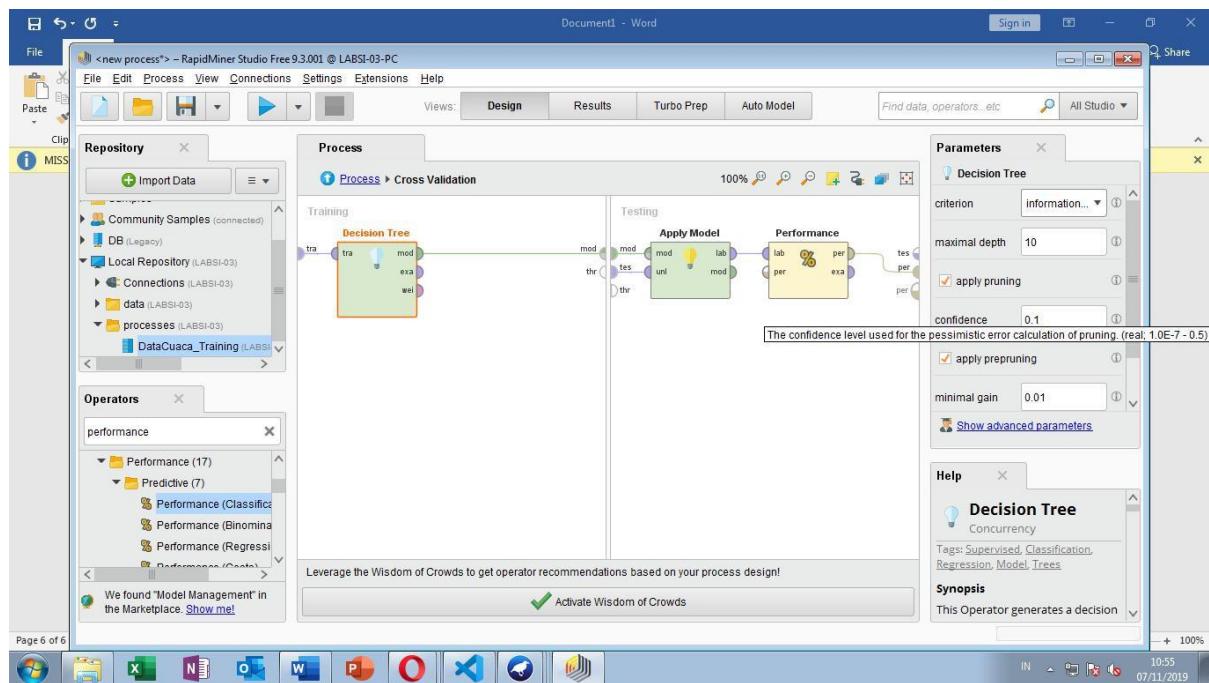
The screenshot shows the RapidMiner interface with the 'Format your columns' dialog open. The dialog title is 'Format your columns.' and it contains a table with 13 rows of data. The columns are labeled: Cuaca (polynomial), Suhu (integer), Kelembaban_u... (integer), Berangin (polynomial), and Bermain_Tenis (binominal/label). The data rows show various weather conditions (Cerah, Mendung, Hujan) and their corresponding values for temperature, humidity, wind, and whether people play tennis. A green checkmark at the bottom right indicates 'no problems.'

Cuaca (polynomial)	Suhu (integer)	Kelembaban_u... (integer)	Berangin (polynomial)	Bermain_Tenis (binominal/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

2. Process View

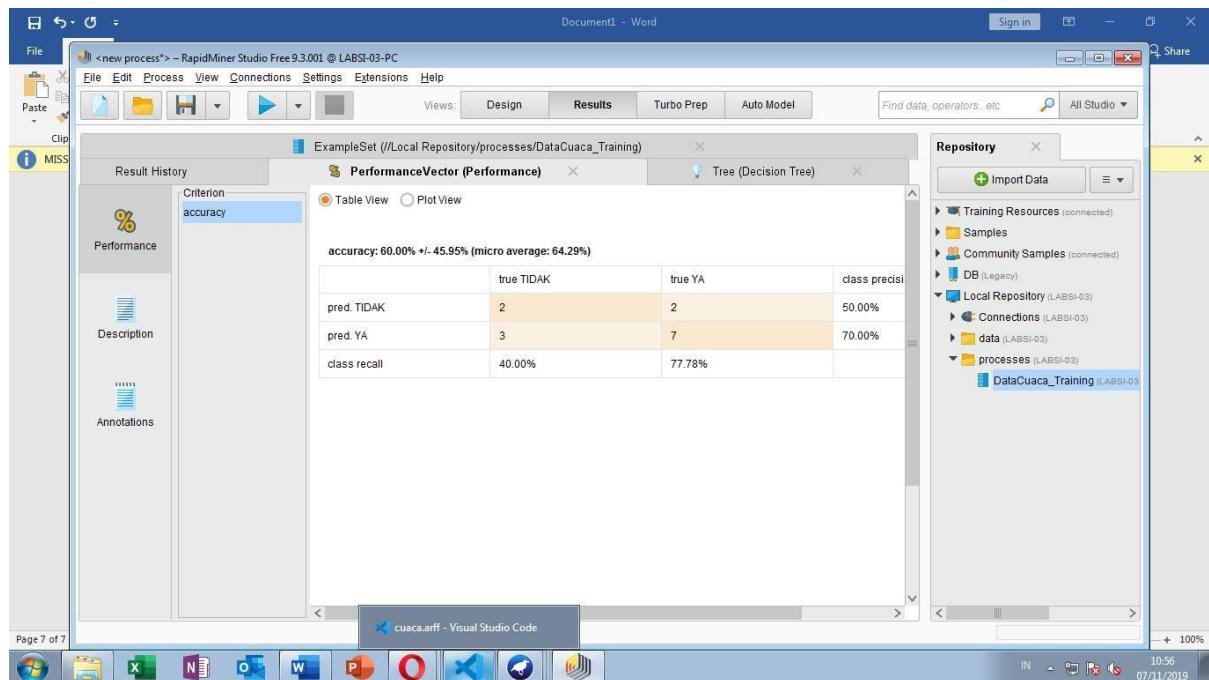
The screenshot shows the RapidMiner process view. A process named 'DataCuaca_Training' is displayed. The process starts with a 'Retrieve DataCuaca_Training' operator, which has an 'out' port connected to the 'inp' port of a 'Cross Validation' operator. The 'Cross Validation' operator has several output ports: 'exa', 'mod', 'exa', 'tes', 'per', and 'par', each leading to a 'res' port. The process view also includes a sidebar with operators like 'Validation', 'Cross Validation', 'Utility', and 'Macros'. A message at the bottom encourages activating the 'Wisdom of Crowds' feature.

3. Process Validation

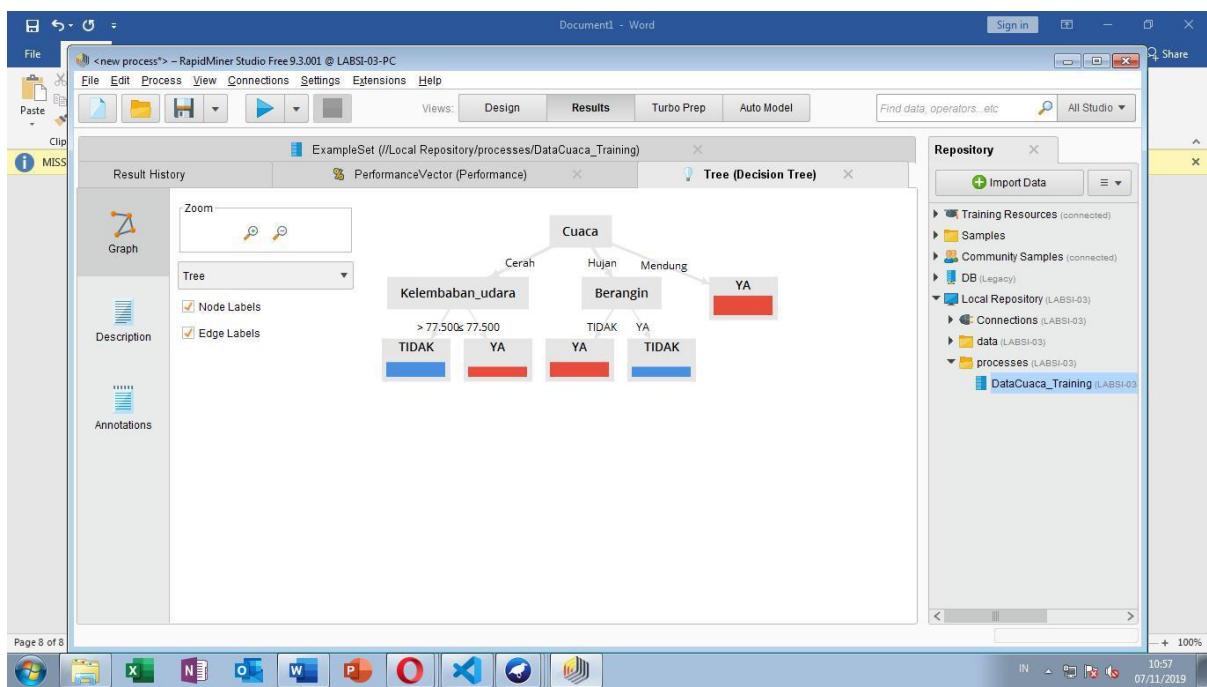


4. Hasil proses klasifikasi pohon keputusan.

a) PerformanceVector



b) Tree (Decision Tree)



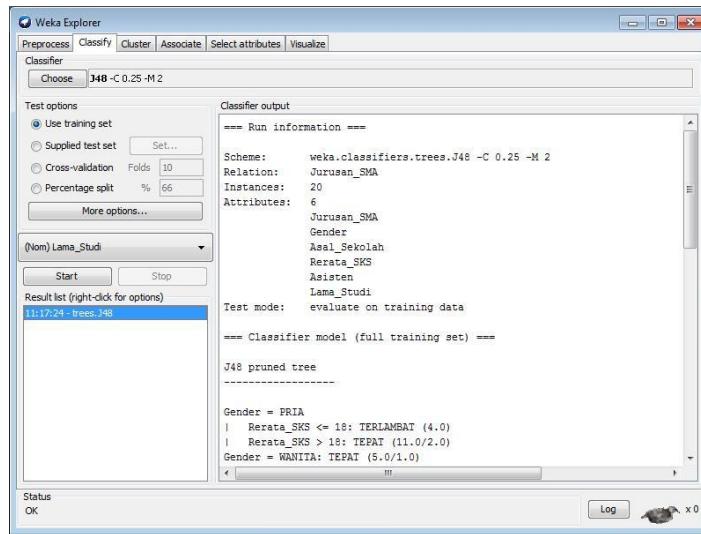
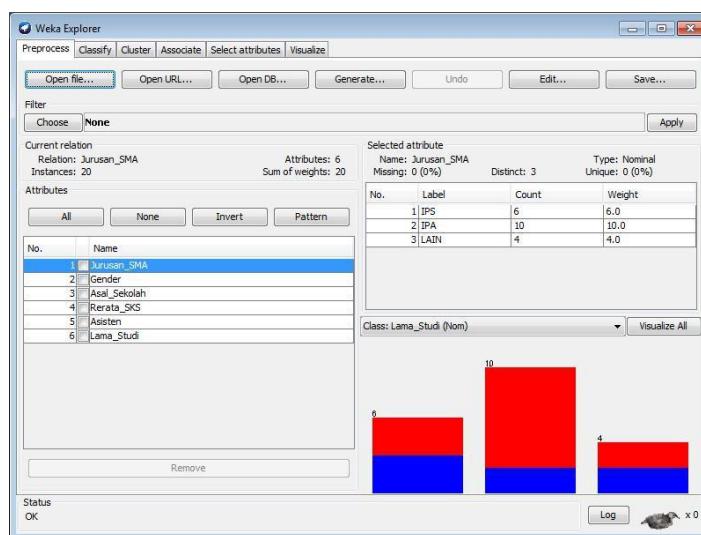
Tugas

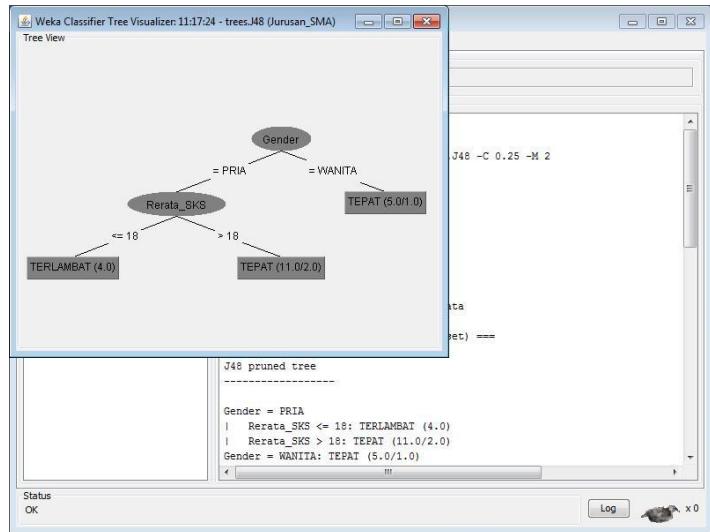
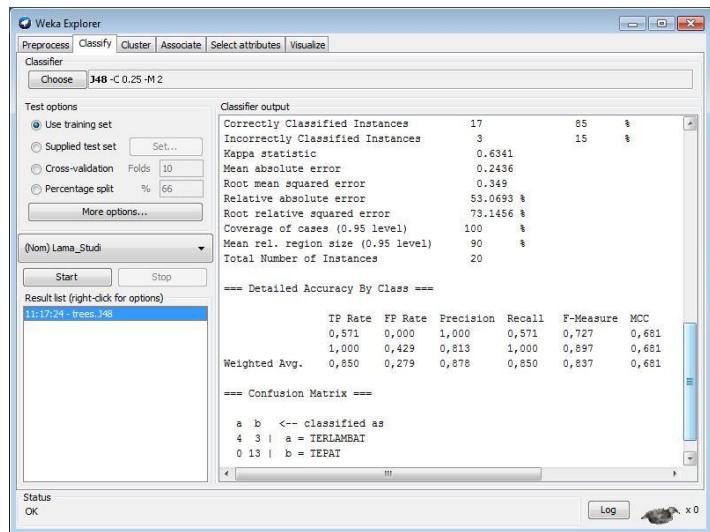
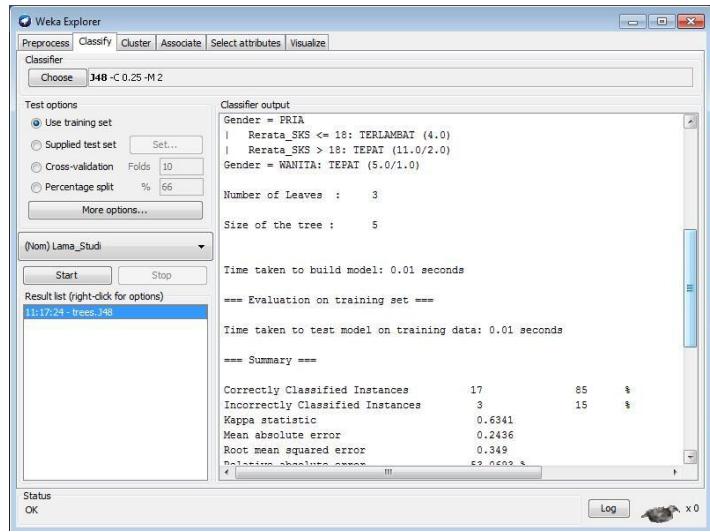
- Nilai kelas atribut Bermain_Tenis.

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

- Gunakan File jurusan_SMA.arff

a.) Membuat Pohon Keputusan menggunakan file jurusan_SMA.arff





b.) Carilah nilai-nilai parameter berikut:

- i. Jumlah simpul daun pada pohon keputusan = 3
- ii. Jumlah simpul keseluruhan pada pohon keputusan = 5
- iii. Waktu yang dibutuhkan untuk proses pelatihan = 0.01 detik
- iv. Tingkat ketepatan klasifikasi = 85 %
- v. Tingkat ketidaktepatan klasifikasi = 15 %

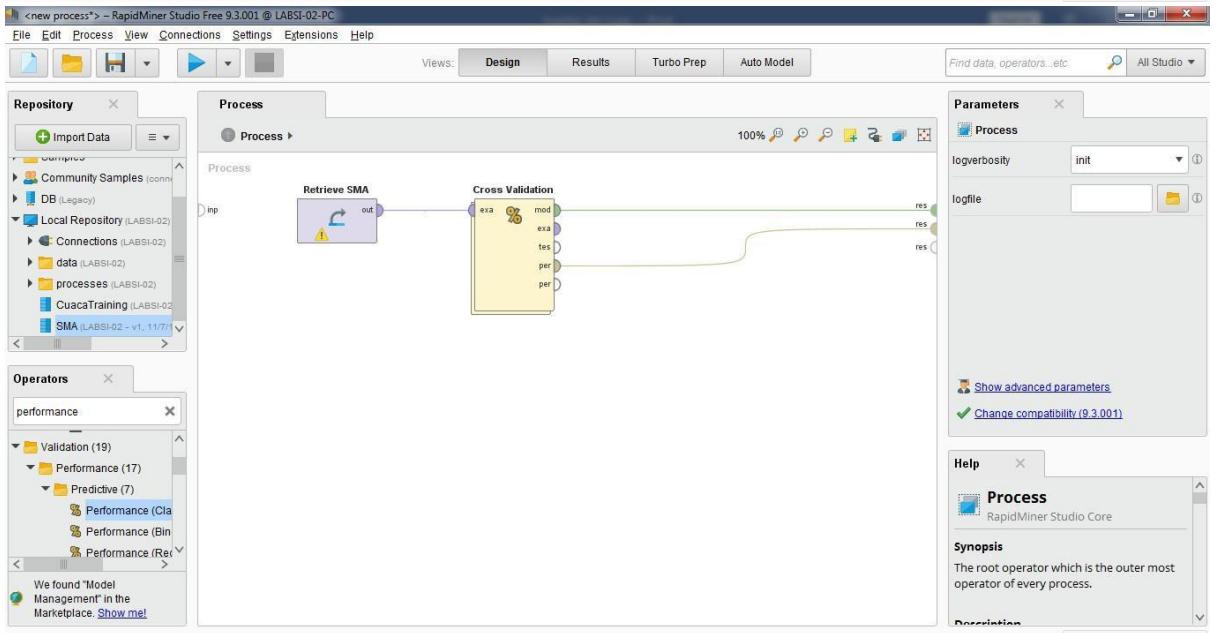
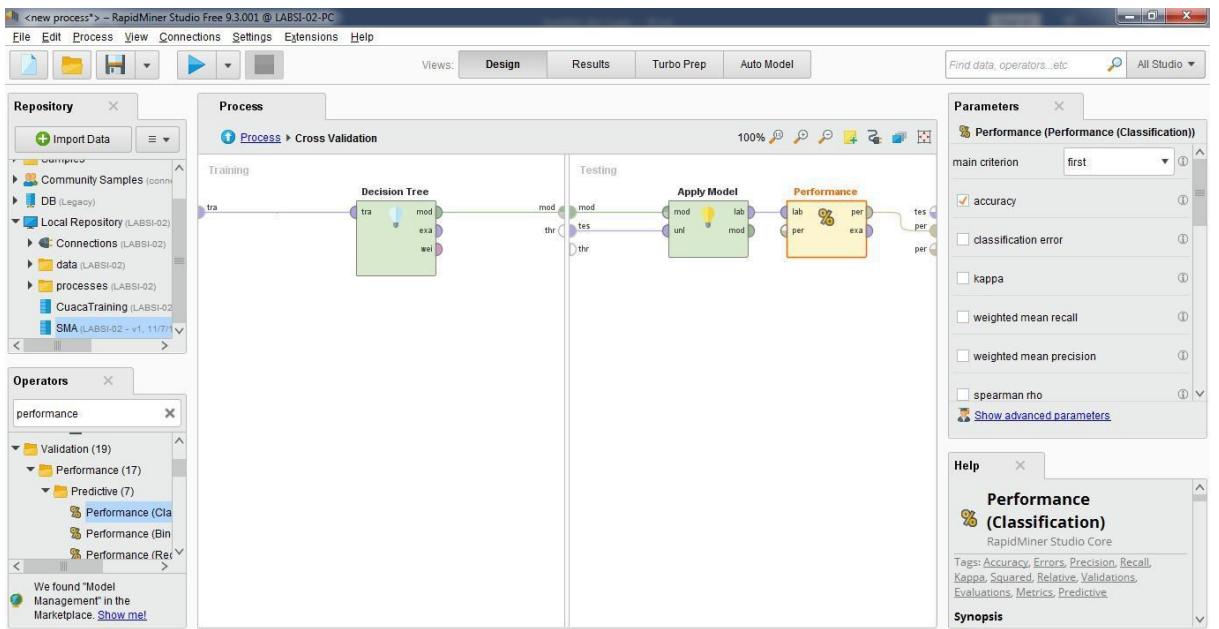
3. Menggunakan file Excel jurusan_SMA.

The screenshot shows the RapidMiner Studio interface with the 'Format your columns' dialog box open. The dialog box is titled 'Format your columns.' and contains a table with 13 rows and 7 columns. The columns are labeled: Jurusan_SMA, Gender, Asal_Sekolah, Rerata_SKS, Asisten, and Lama_Studi. The last column is labeled 'binomial label'. The data in the table is as follows:

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Label
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	SURAKARTA	TIDAK	TEPAT

The dialog box also includes a checkbox for 'Replace errors with missing values' and a button for 'no problems.'

a. Pohon Keputusan



RapidMiner Studio Free 9.3.001 @ LABSI-02-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators...etc All Studio

Repository

- Training Resources (connected)
- Samples
- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-02)
 - Connections (LABSI-02)
 - data (LABSI-02)
 - processes (LABSI-02)
 - CuacaTraining (LABSI-02 - v1, 11/7/19 10:44 AM - 569 bytes)
 - SMA (LABSI-02 - v1, 11/7/19 11:24 AM - 669 bytes)

PerformanceVector (Performance)

Criterion: accuracy

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%	

Tree (Decision Tree)

Zoom

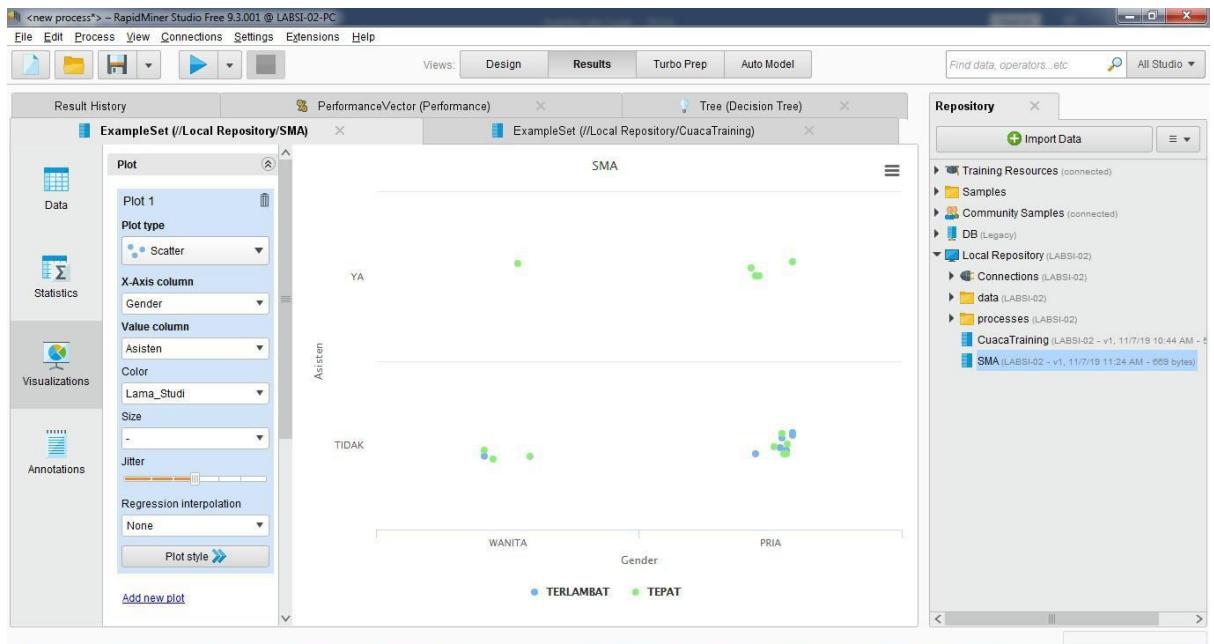
Tree

Node Labels Edge Labels

```

graph TD
    Asisten[Asisten] -- TIDAK --> Rerata_SKS[Rerata_SKS]
    Asisten -- YA --> TEPAT[TEPAT]
    Rerata_SKS -- "≤ 18.500" --> Gender[Gender]
    Rerata_SKS -- "> 18.500" --> TEPAT[TEPAT]
    Gender -- PRIA --> TERLAMBAT[TERLAMBAT]
    Gender -- WANITA --> TEPAT[TEPAT]
  
```

b. Perspektif Plot View



4. Klasifikasi yang terbentuk yaitu :

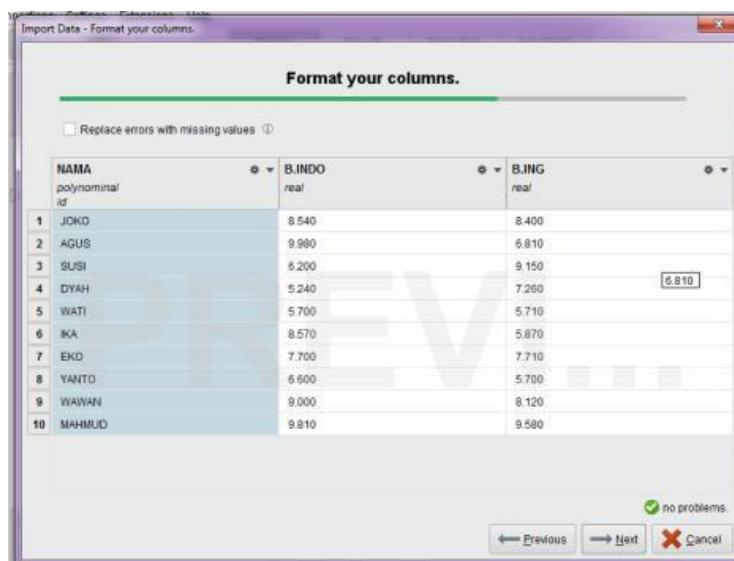
18. Seseorang akan lama_studi (TERLAMBAT) jika kondisi sebagai berikut:
 - Gender = PRIA, Rerata_SKS <= 18 (nilai atribut lain diabaikan)
19. Seseorang akan lama_studi (TERLAMBAT) jika kondisi sebagai berikut:
 - Gender = WANITA (nilai atribut lain diabaikan)
 - Gender = PRIA, Rerata_SKS > 18 (nilai atribut lain diabaikan)

MODUL 10

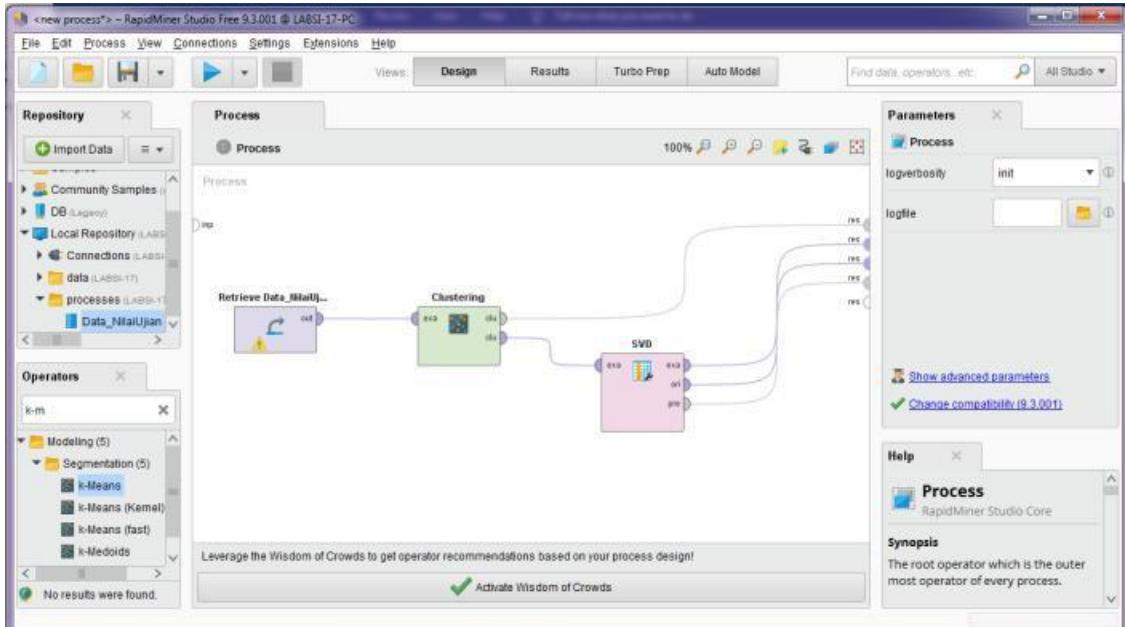
Kegiatan

1. Tabel Data Nilai Siswa

NO_SISWA	NAMA	B.INDO	B.ING
S-101	JOKO	8,54	8,4
S-102	AGUS	9,98	6,81
S-103	SUSI	6,2	9,15
S-104	DYAH	5,24	7,26
S-105	WATI	5,7	5,71
S-106	IKA	8,57	5,87
S-107	EKO	7,7	7,71
S-108	YANTO	6,6	5,7
S-109	WAWAN	9	8,12
S-110	MAHMUD	9,81	9,58



2. Process



3. Hasil Proses Clustering dengan Algoritma K-Means

a) SVD (Singular Value Decomposition)

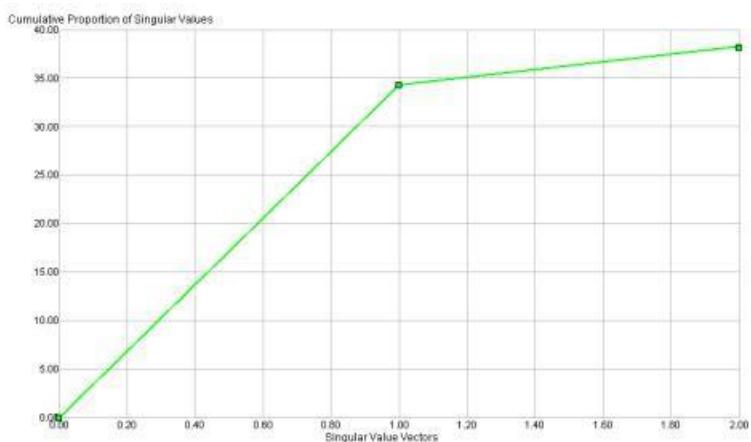
i. Nilai Eigenvalue

Component	Singular Value	Proportion of Singular V...	Cumulative Singular Val...	Cumulative Proportion o...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

ii. Nilai SVD Vectors

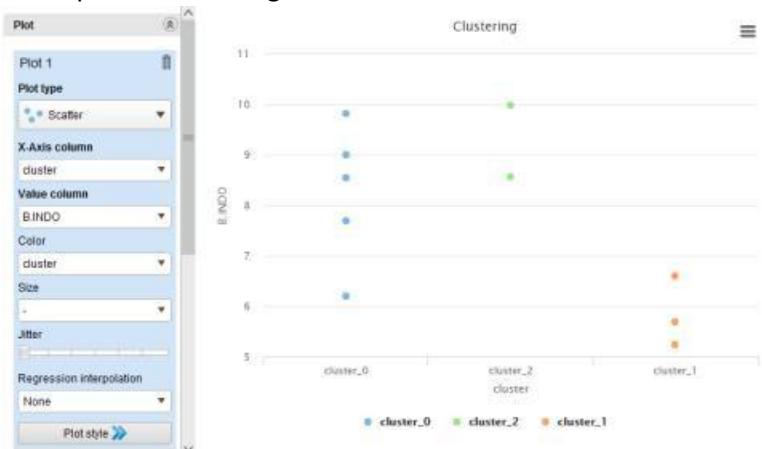
Attribute	SVD Vector 1
B.INDO	0.723
B.ING	0.690

iii. Nilai Cumulative Variance



b) ExampleSet (K-Means)

I. Kelompok Siswa bidang B.Indonesia



II. Kelompok Siswa bidang B. Inggris



c) 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

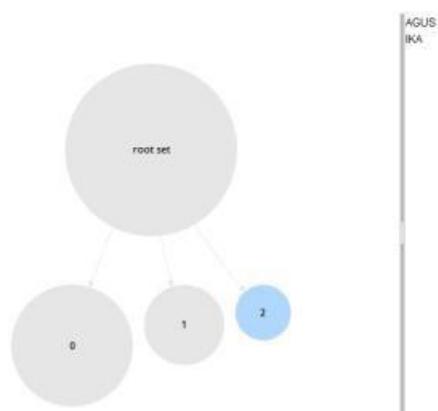
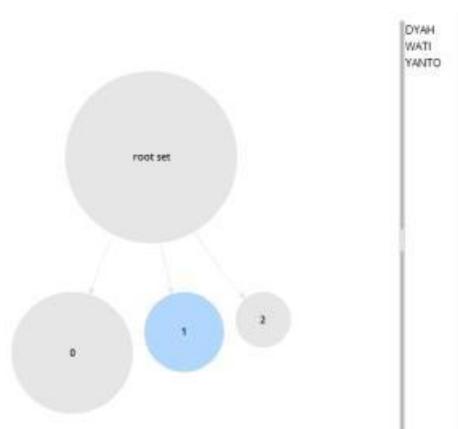
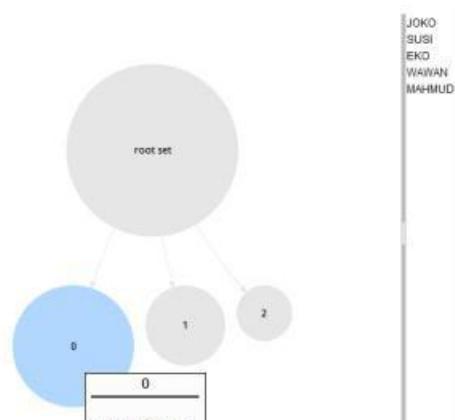
d) Cluster Model (Clustering)

- Description

Cluster Model

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

- Graph

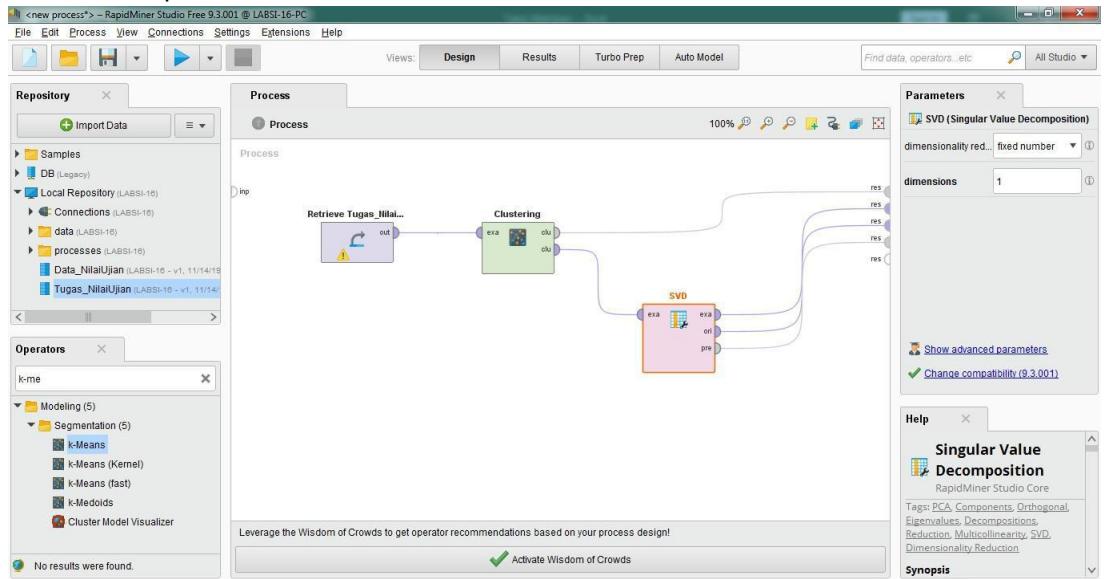


Tugas

1. Tabel Data Nilai Ujian

	A	B	C	D	E	F	G
1	NO_SISWA	NAMA	B.IND	B.ING	MTK	IPA	
2	S_101	JOKO	9,95	5,40	7,41	8,32	
3	S_102	AGUS	9,37	5,91	8,56	6,43	
4	S_103	SUSI	8,77	5,14	5,38	5,94	
5	S_104	DYAH	7,62	6,02	5,38	9,02	
6	S_105	WATI	9,61	6,45	7,94	8,39	
7	S_106	IKA	7,72	7,22	6,24	8,89	
8	S_107	EKO	5,07	8,25	5,73	7,81	
9	S_108	YANTO	7,42	9,97	8,87	9,70	
10	S_109	WAWAN	5,50	7,04	8,91	8,33	
11	S_110	MAHMUD	8,82	9,13	8,51	9,64	
12	S_111	BUDI	9,17	8,05	5,74	5,55	
13	S_112	SANTI	6,77	5,11	5,52	6,80	
14	S_113	DIAN	7,85	9,90	9,37	5,10	
15	S_114	DANI	8,60	6,21	7,38	7,83	
16	S_115	AHMAD	6,31	8,64	9,23	6,14	
17	S_116	BAYU	8,86	6,67	8,81	5,35	
18	S_117	RISA	6,18	6,86	6,42	7,78	
19	S_118	RANI	9,04	8,00	9,82	7,16	
20	S_119	YANI	5,20	5,24	5,59	8,50	
21	S_120	RATIH	7,96	5,64	8,72	6,15	
22	S_121	INDAH	7,45	6,29	8,15	9,06	
23	S_122	JONO	6,76	8,89	6,39	8,09	
24	S_123	SARAH	8,51	7,64	6,95	6,06	
25	S_124	RAMA	6,42	6,07	7,63	5,89	
26	S_125	BAMBANG	9,06	7,71	8,52	7,41	
27	S_126	HADI	5,41	5,65	6,66	5,48	
28	S_127	NANA	5,81	8,43	7,01	6,43	
29	S_128	FEBRI	7,56	7,84	8,25	7,67	
30	S_129	DENI	9,51	9,82	7,64	6,70	

2. Process di RapidMiner



3. Hasil Proses Clustering dengan Algoritma K-Means

a) SVD (Singular Value Decomposition)

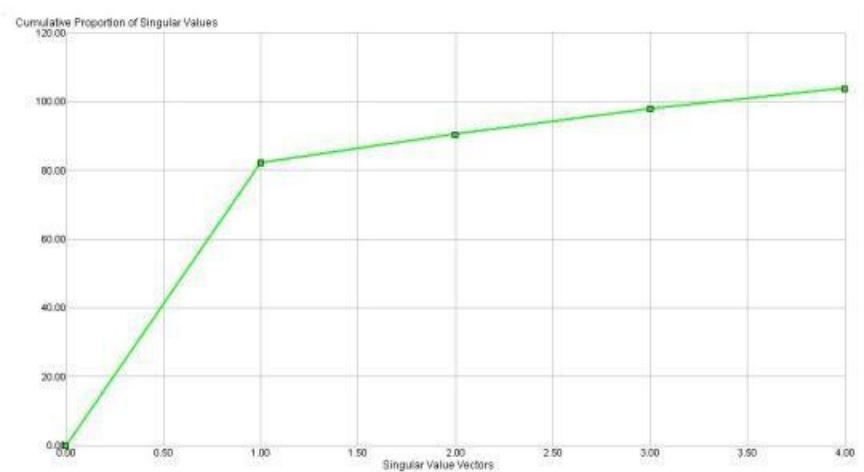
i. Nilai Eigenvalue

Component	Singular Value	Proportion of Singular Values	Cumulative Singular Values	Cumulative Proportion of Sin...
SVD 1	82.312	0.792	82.312	0.792
SVD 2	8.172	0.079	90.484	0.871
SVD 3	7.462	0.072	97.947	0.943
SVD 4	5.968	0.057	103.914	1.000

ii. Nilai SVD Vectors

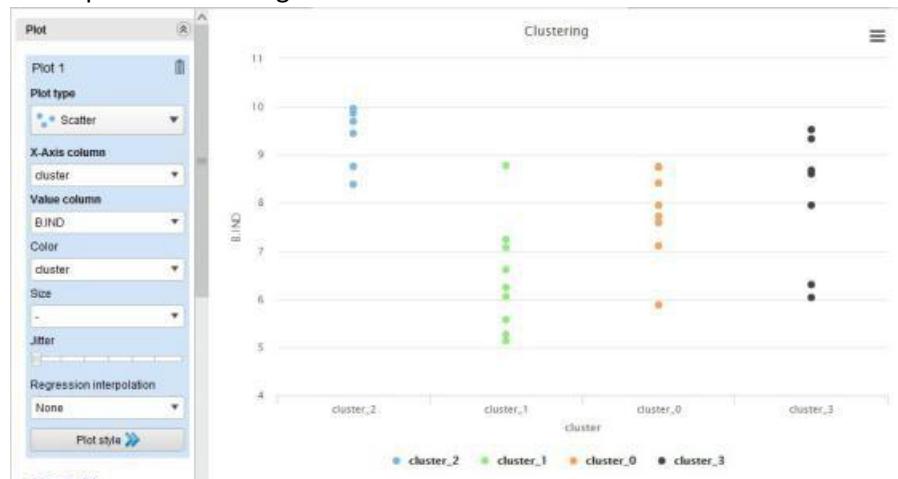
Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.IND	0.520	-0.111	0.489
B.ING	0.470	-0.770	-0.075
MTK	0.524	0.581	0.325
IPA	0.483	0.238	-0.805

iii. Nilai Cumulative Variance

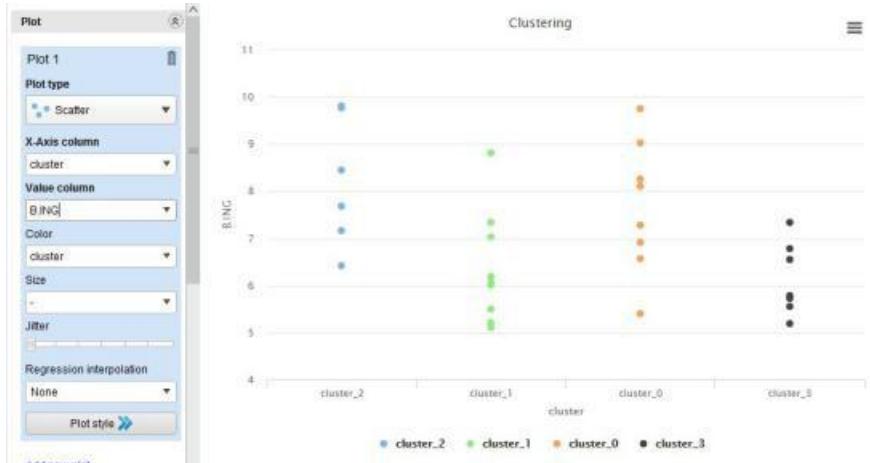


b) ExampleSet (K-Means)

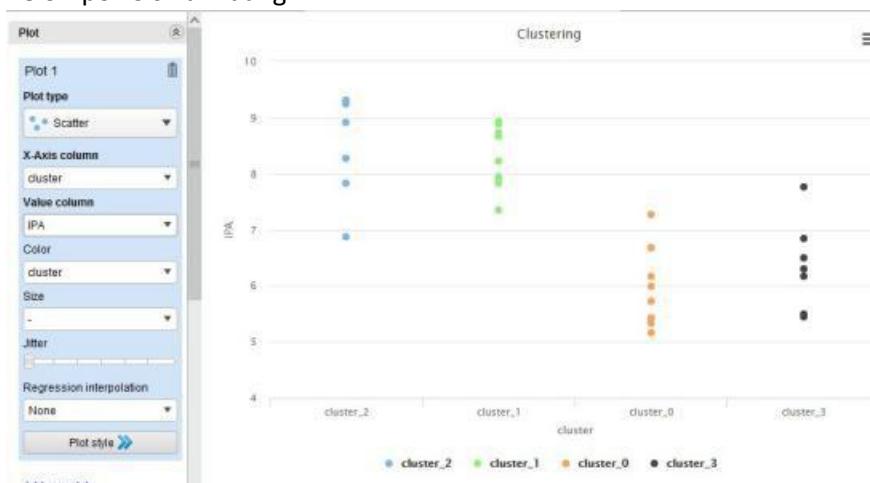
I. Kelompok Siswa Bidang B.Indonesia



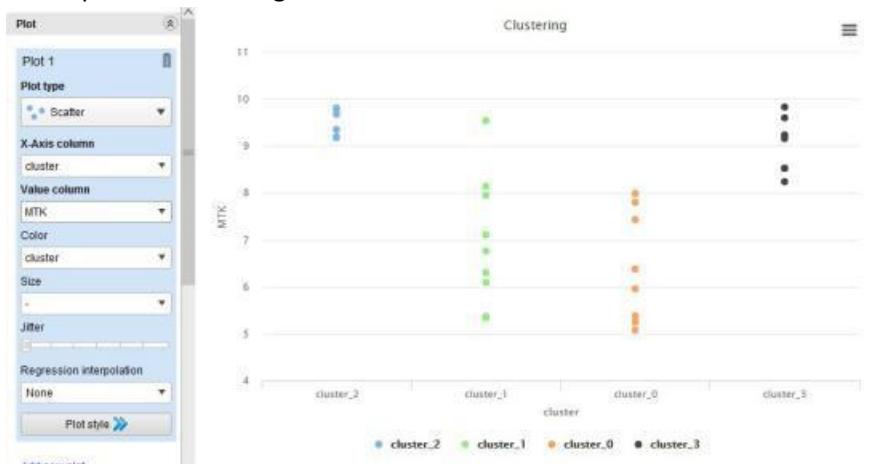
II. Kelompok Siswa Bidang B. Inggris



III. Kelompok Siswa Bidang IPA



IV. Kelompok Siswa Bidang MTK



c) ExampleSet (SVD)

Row No.	NAMA	cluster_↑	BIND	BJNG	MTK	IPA
3	SUSI	cluster_0	8.761	9.029	5.078	5.329
5	WATI	cluster_0	7.729	6.909	6.373	7.272
6	IKA	cluster_0	7.108	8.251	5.954	5.434
7	EKO	cluster_0	5.887	7.280	7.988	5.985
9	WAWAN	cluster_0	8.412	5.404	5.243	5.728
18	RANI	cluster_0	8.742	8.111	7.799	8.174
20	RATIH	cluster_0	7.584	9.748	5.375	6.686
23	SARAH	cluster_0	7.948	6.575	7.434	5.172
2	AGUS	cluster_1	8.766	6.193	6.296	8.662
8	YANTO	cluster_1	5.266	7.035	6.096	8.886
10	MAHMUD	cluster_1	5.568	5.494	6.760	7.929
12	SANTI	cluster_1	7.247	6.046	7.104	7.838
13	DIAN	cluster_1	6.235	8.819	7.954	8.229
16	BAYU	cluster_1	7.061	6.022	8.134	8.938
22	JONO	cluster_1	6.607	7.338	5.336	7.949
24	RAMA	cluster_1	5.135	5.214	5.364	7.357
26	HADI	cluster_1	6.056	5.108	9.539	8.736
1	JOKO	cluster_2	8.384	9.815	9.782	6.875
11	BUDI	cluster_2	9.862	7.676	9.172	7.827
14	DANI	cluster_2	9.699	7.158	9.680	8.909
15	AHMAD	cluster_2	8.748	8.450	9.205	8.268
28	FEBRI	cluster_2	9.453	9.779	9.820	9.317
30	TONI	cluster_2	9.960	6.408	9.359	9.251
4	DYAH	cluster_3	9.338	6.785	9.153	5.438
17	RISA	cluster_3	8.679	5.779	8.231	7.768
19	YANI	cluster_3	9.516	6.542	9.608	6.306
21	INDAH	cluster_3	7.959	5.190	8.518	6.843
25	BAMBANG	cluster_3	6.303	5.548	8.530	5.503
27	NANA	cluster_3	8.606	5.720	9.228	6.506
29	DENI	cluster_3	6.038	7.337	9.835	6.172

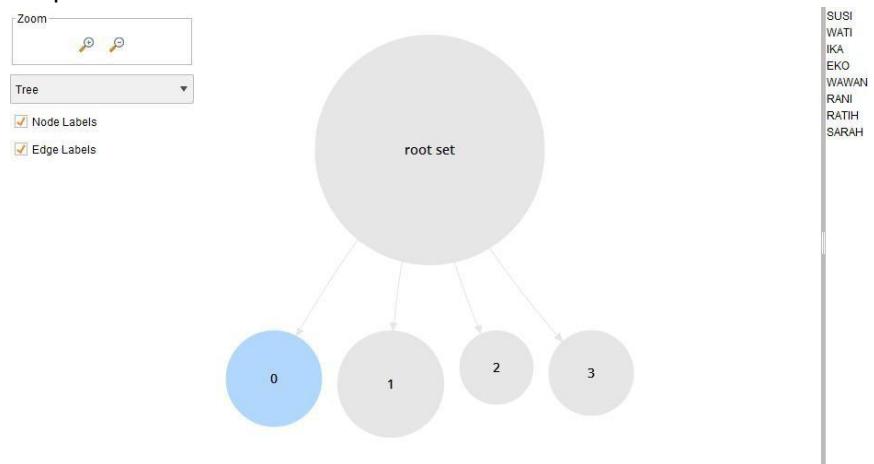
d) Cluster Model (Clustering)

- Description

Cluster Model

Cluster 0: 8 items
 Cluster 1: 9 items
 Cluster 2: 6 items
 Cluster 3: 7 items
 Total number of items: 30

- Graph





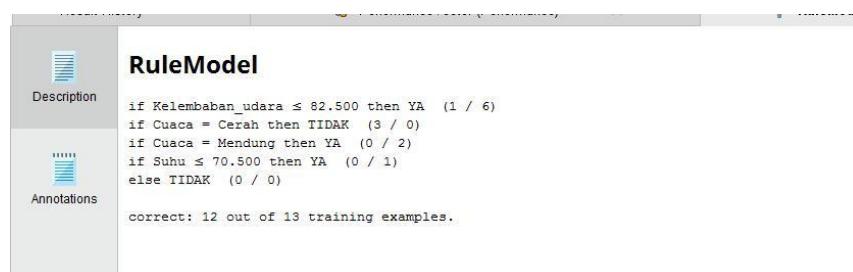
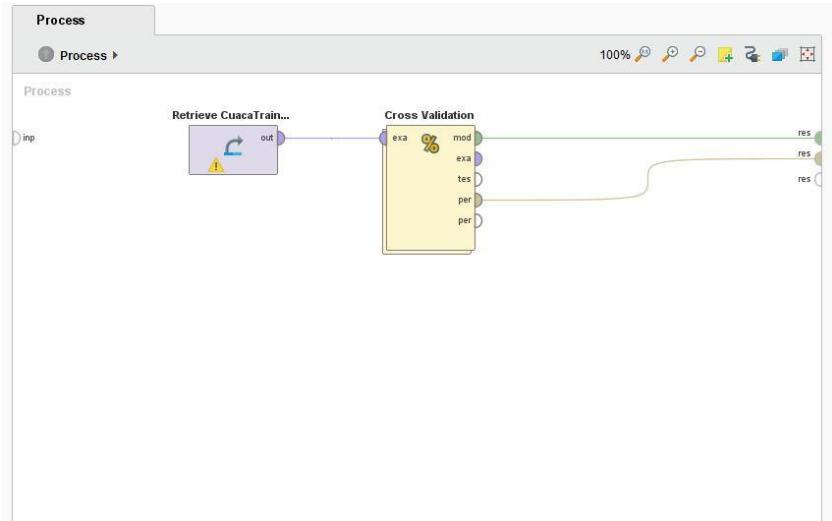
Nama siswa yang terdapat dalam cluster :

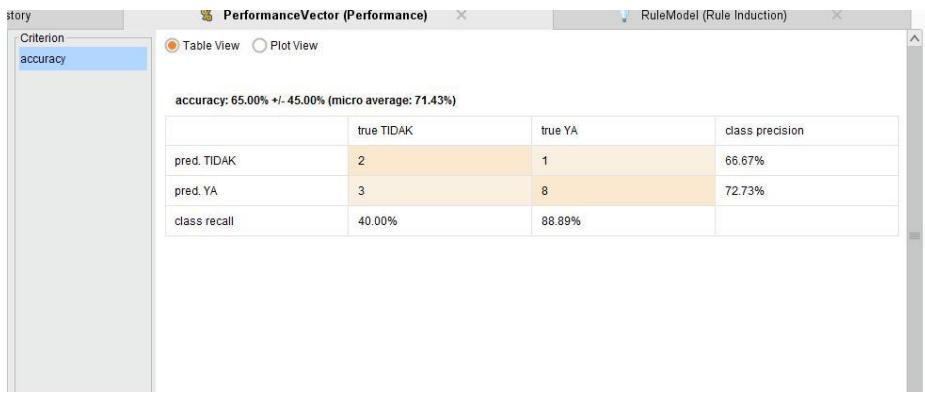
- Cluster 0 : Susi, Wati, Ika, Eko, Wawan, Rani, Ratih, Sarah
- Cluster 1 : Agus, Yanto, Mahmud, Santi, Dian, Bayu, Jono, Rama, Hadi
- Cluster 2 : Joko, Budi, Dani, Ahmad, Febri, Toni
- Cluster 3 : Dyah, Risa, Yani, Indah, Bambang, Nana, Deni

MODUL 11

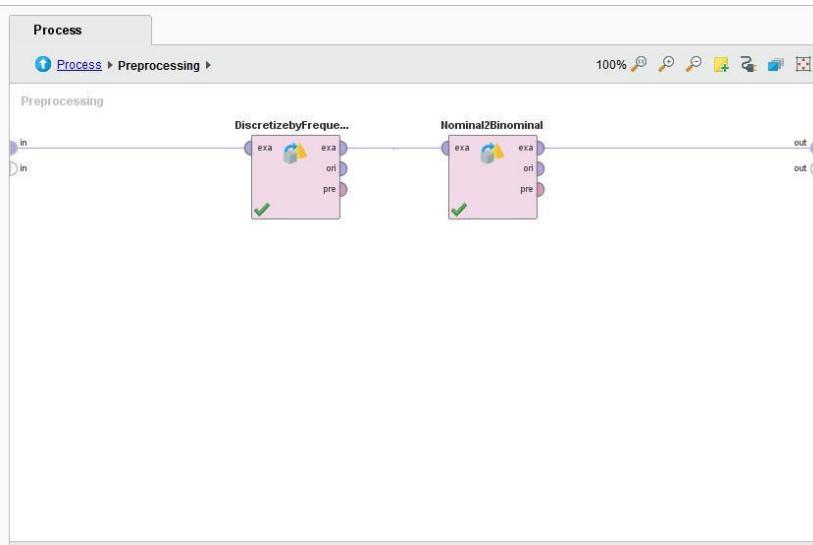
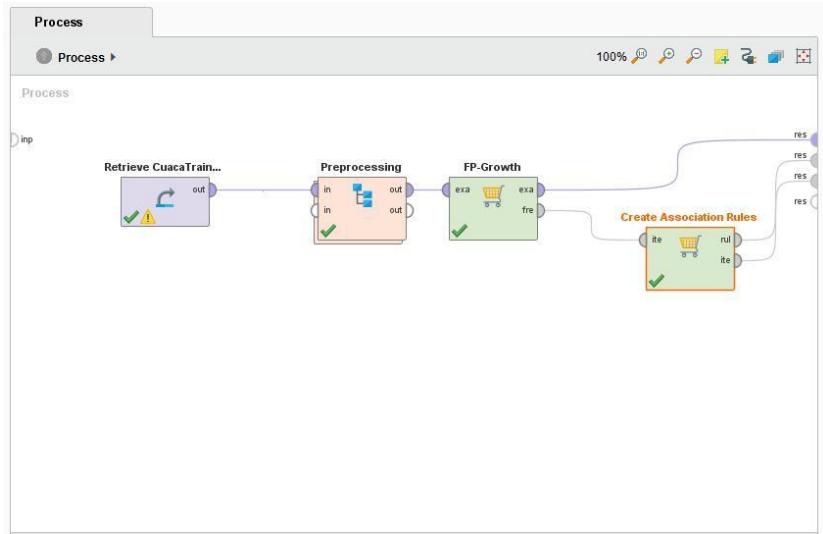
Kegiatan

1. Induksi Aturan Cuaca





2. Aturan Asosiasi Data Cuaca



a) Frequent Item Set

The screenshot shows the RapidMiner interface with the 'FrequentItemSets (FP-Growth)' tab selected. The results table displays frequent item sets of size 1 to 4 with their support values and item descriptions.

Size	Support	Item 1	Item 2	Item 3	Item 4
1	0.500	Kelambaban_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	Kelambaban_udara	Berangin		
2	0.214	Kelambaban_udara	Suhu		
2	0.214	Kelambaban_udara	Cuaca = Cerah		
2	0.143	Kelambaban_udara	Cuaca = Hujan		
2	0.143	Kelambaban_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		

The screenshot shows the RapidMiner interface with the 'FrequentItemSets (FP-Growth)' tab selected. The results table displays frequent item sets of size 1 to 4 with their support values and item descriptions, including additional items like 'Cuaca = Cerah' and 'Cuaca = Hujan'.

Size	Support	Item 1	Item 2	Item 3	Item 4
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		
2	0.071	Suhu	Cuaca = Hujan		
2	0.143	Suhu	Cuaca = Mendung		
3	0.071	Kelambaban_udara	Berangin	Suhu	
3	0.071	Kelambaban_udara	Berangin	Cusca = Cerah	
3	0.071	Kelambaban_udara	Berangin	Cuaca = Hujan	
3	0.071	Kelambaban_udara	Berangin	Cuaca = Mendung	
3	0.143	Kelambaban_udara	Suhu	Cuaca = Cerah	
3	0.071	Kelambaban_udara	Suhu	Cuaca = Mendung	
3	0.143	Berangin	Suhu	Cuaca = Cerah	
4	0.071	Kelambaban_udara	Berangin	Suhu	Cuaca = Cerah

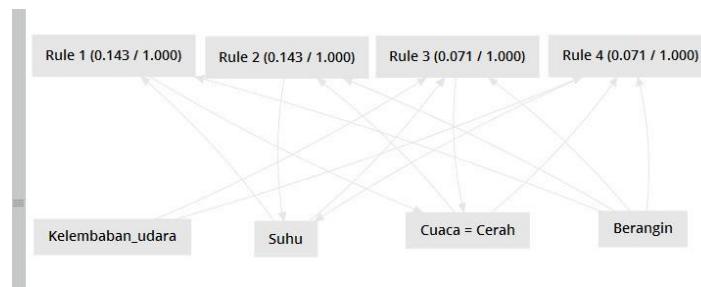
b) Association Rules

i. Table View

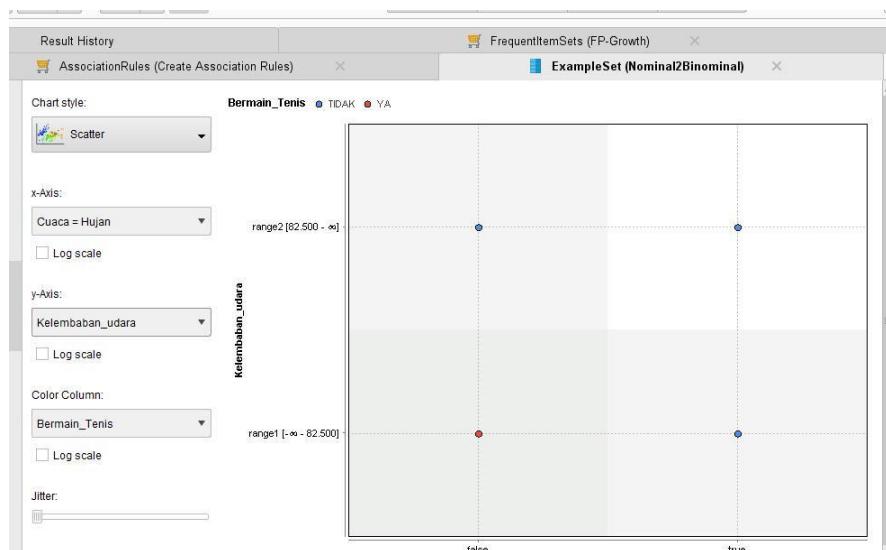
No.	Premises	Conclusion	Support
1	Berangin, Suhu	Cuaca = Cerah	0.143
2	Berangin, Cuaca = Cerah	Suhu	0.143
3	Kelambaban_udara, Berangin, Suhu	Cuaca = Cerah	0.071
4	Kelambaban_udara, Berangin, Cuaca = Cerah	Suhu	0.071

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

ii. Graph View

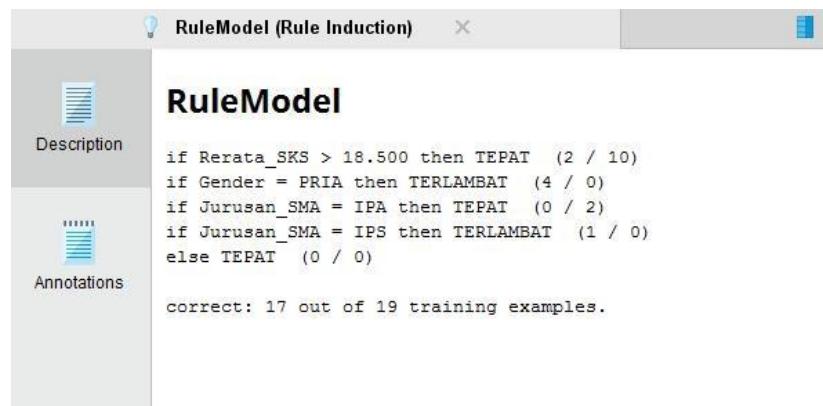
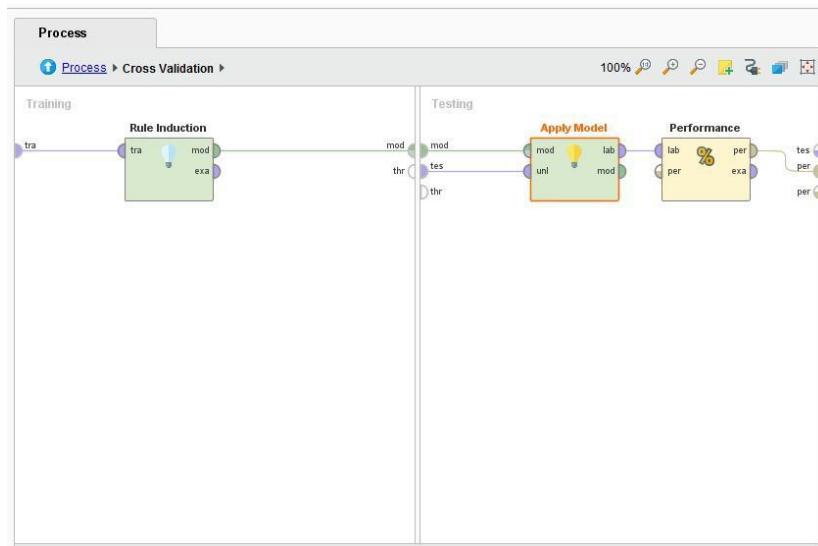
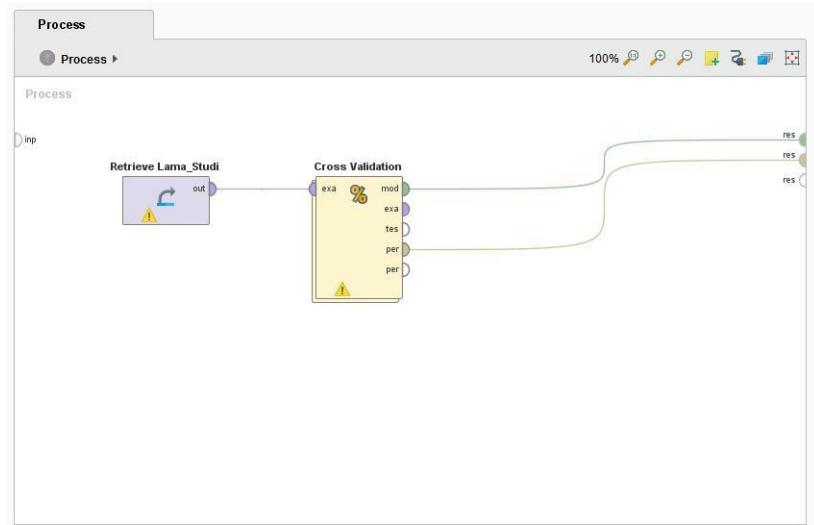


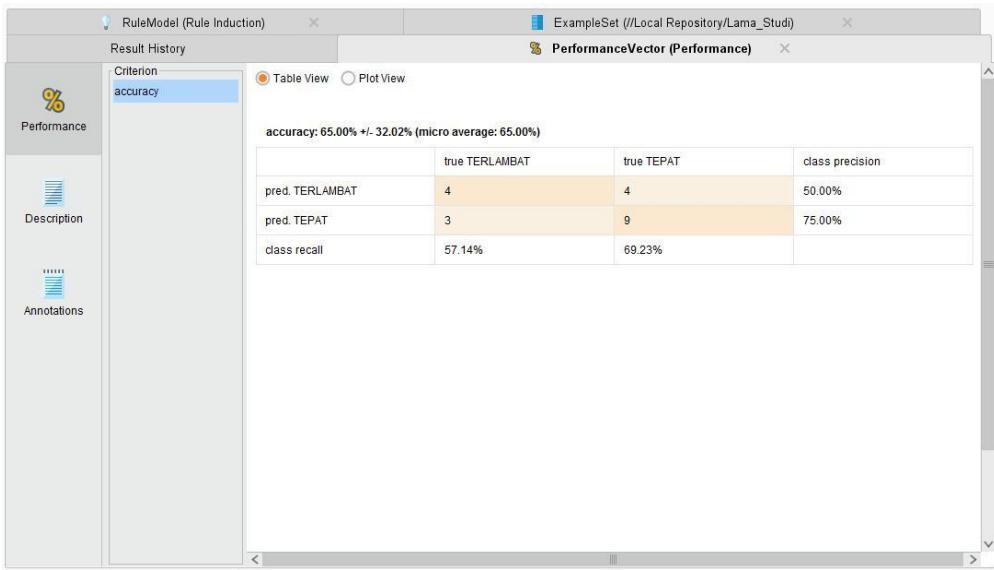
c) Example set



Tugas

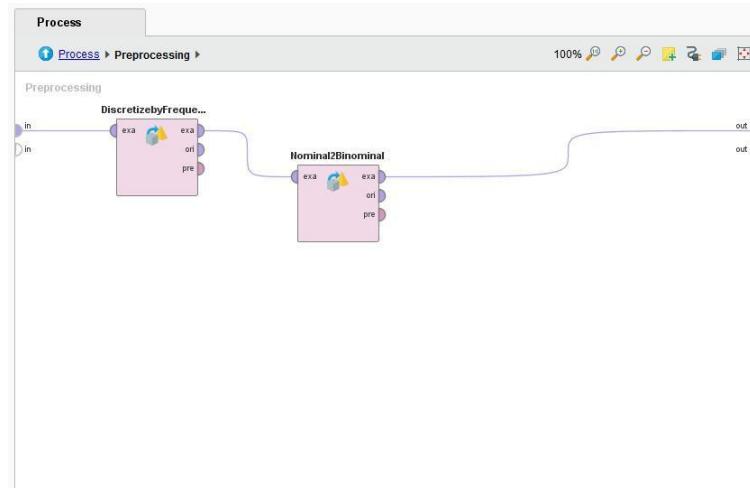
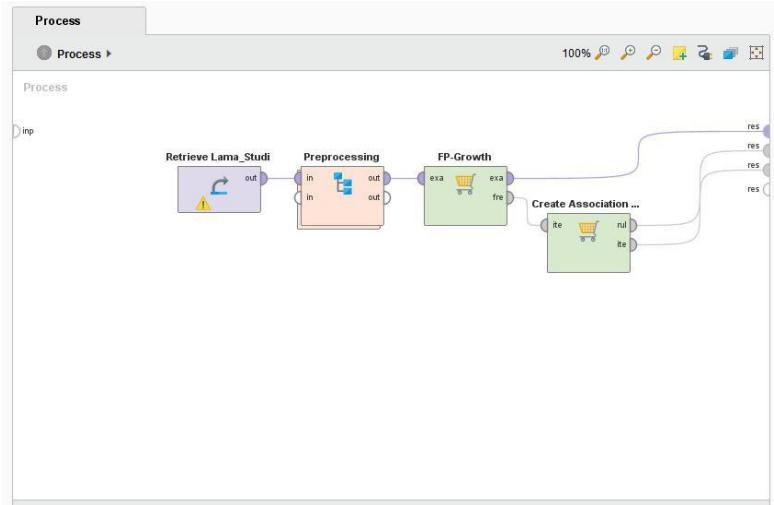
1. Rule Model dan Nilai Performance Vector





2. Nilai :

a) $number\ of\ bins = 2$



Views: Design Results Turbo Prep Auto Model

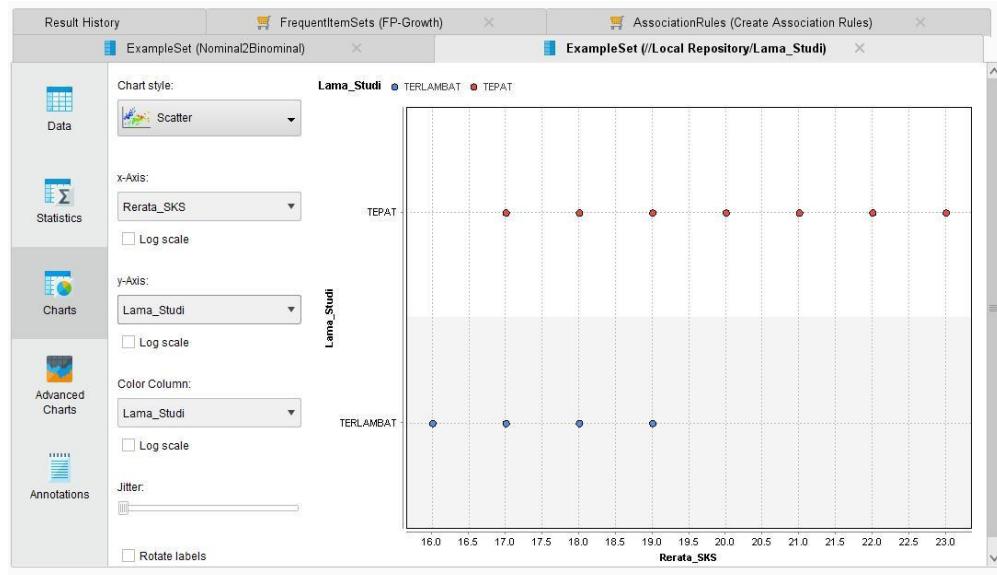
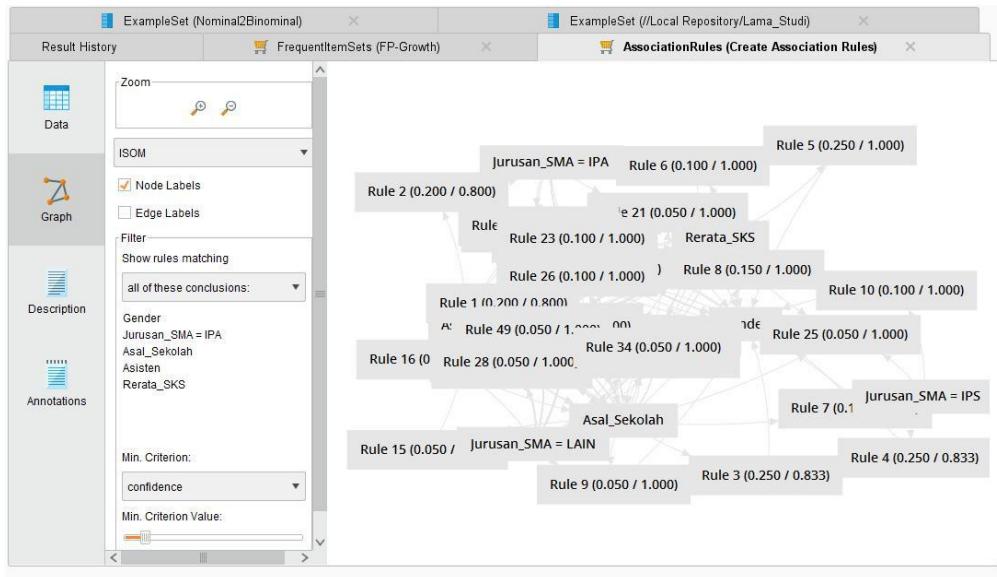
Result History FrequentItemSets (FP.Growth) AssociationRules (Create Association Rules)

Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
1	0.750	Gender				
1	0.500	Jurusan_SMA = ...				
1	0.300	Asal_Sekolah				
1	0.300	Jurusan_SMA = ...				
1	0.250	Asisten				
1	0.250	Rerata_SKS				
1	0.200	Jurusan_SMA = ...				
2	0.350	Gender	Jurusan_SMA = ...			
2	0.250	Gender	Asal_Sekolah			
2	0.250	Gender	Jurusan_SMA = ...			
2	0.200	Gender	Asisten			
2	0.250	Gender	Rerata_SKS			
2	0.150	Gender	Jurusan_SMA = ...			
2	0.150	Jurusan_SMA = ...	Asal_Sekolah			
2	0.200	Jurusan_SMA = ...	Asisten			

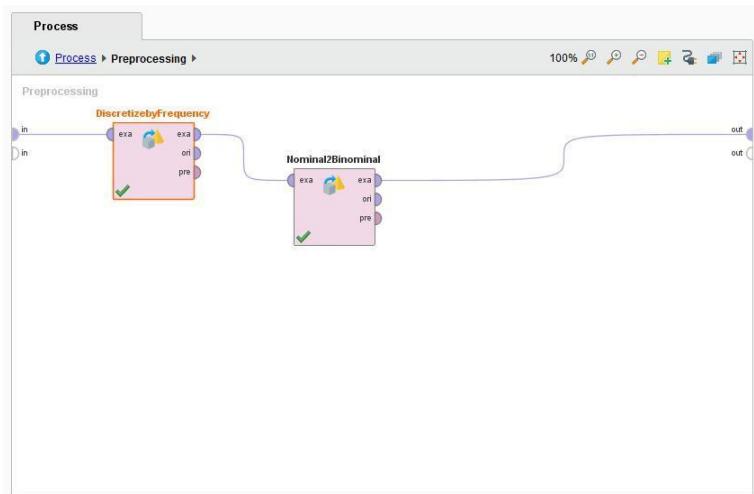
Views: Design Results Turbo Prep Auto Model

Result History FrequentItemSets (FP.Growth) AssociationRules (Create Association Rules)

No.	Premises	Conclusion	Support
17	Asal_Sekolah, Jurusan_SMA = LAIN	Rerata_SKS	0.050
19	Asisten, Jurusan_SMA = LAIN	Rerata_SKS	0.050
26	Gender, Asal_Sekolah, Asisten	Rerata_SKS	0.100
33	Asal_Sekolah, Jurusan_SMA = LAIN	Gender, Rerata_SKS	0.050
34	Gender, Asal_Sekolah, Jurusan_SMA = LAIN	Rerata_SKS	0.050
38	Asisten, Jurusan_SMA = LAIN	Gender, Rerata_SKS	0.050
39	Gender, Asisten, Jurusan_SMA = LAIN	Rerata_SKS	0.050
44	Asal_Sekolah, Jurusan_SMA = LAIN	Asisten, Rerata_SKS	0.050
45	Asisten, Jurusan_SMA = LAIN	Asal_Sekolah, Rerata_SKS	0.050
46	Asal_Sekolah, Asisten, Jurusan_SMA = LAIN	Rerata_SKS	0.050
50	Gender, Jurusan_SMA = IPA, Asal_Sekolah, Asisten	Rerata_SKS	0.050
54	Asal_Sekolah, Jurusan_SMA = LAIN	Gender, Asisten, Rerata_SKS	0.050
55	Gender, Asal_Sekolah, Jurusan_SMA = LAIN	Asisten, Rerata_SKS	0.050
56	Asisten, Jurusan_SMA = LAIN	Gender, Asal_Sekolah, Rerata_SKS	0.050
57	Gender, Asisten, Jurusan_SMA = LAIN	Asal_Sekolah, Rerata_SKS	0.050



b) *number of bins = 3*



ExampleSet (Nominal2Binomial) ExampleSet (/Local Repository/Lama_Studi)

Result History FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules)

Data

Size	Support	Item 1	Item 2	Item 3	Item 4	Item 5
1	0.750	Gender				
1	0.500	Jurusan_SMA = ...				
1	0.400	Rerata_SKS = r...				
1	0.350	Rerata_SKS = r...				
1	0.300	Asal_Sekolah				
1	0.300	Jurusan_SMA = ...				
1	0.250	Asisten				
1	0.250	Rerata_SKS = r...				
1	0.200	Jurusan_SMA = ...				
2	0.350	Gender	Jurusan_SMA = ...			
2	0.200	Gender	Rerata_SKS = r...			
2	0.300	Gender	Rerata_SKS = r...			
2	0.250	Gender	Asal_Sekolah			
2	0.250	Gender	Jurusan_SMA = ...			
2	0.200	Gender	Asisten			

Annotations

Min. Size: 1 Max. Size: 5
Contains Item:
Update View

ExampleSet (Nominal2Binomial) ExampleSet (/Local Repository/Lama_Studi)

Result History FrequentItemSets (FP-Growth) AssociationRules (Create Association Rules)

Data

Show rules matching all of these conclusions:

- Gender
- Jurusan_SMA = IPA
- Asal_Sekolah
- Asisten
- Rerata_SKS = range3 [19.500 - ∞]

The item Rerata_SKS = range3 [19.500 - ∞]

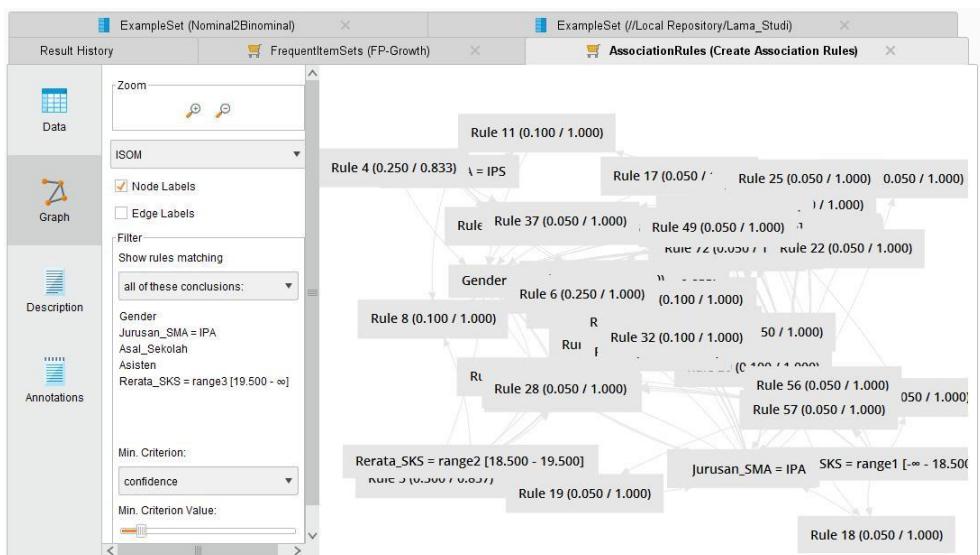
Min. Criterion: confidence

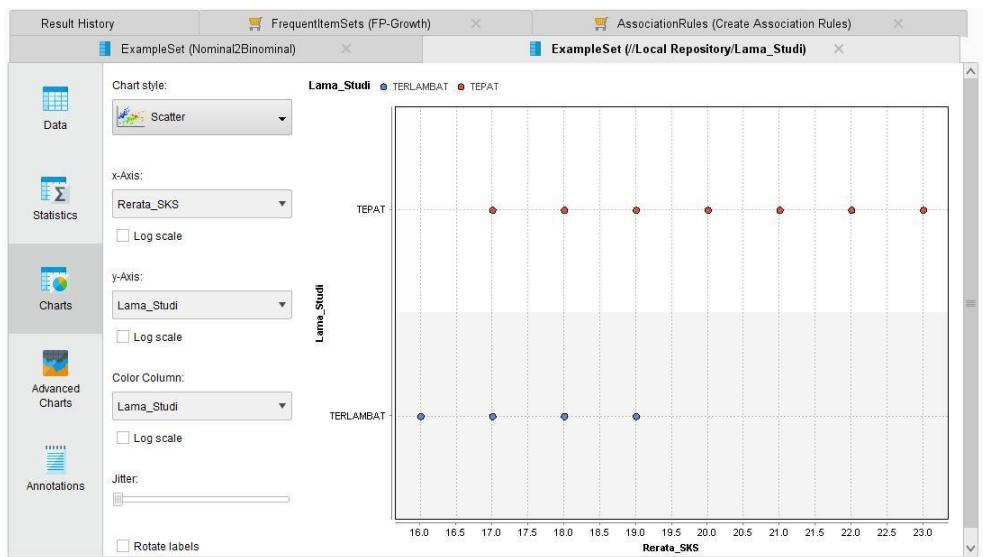
Min. Criterion Value: 0.05

No.	Premises	Conclusion	Support
3	Asal_Sekolah	Gender	0.250
4	Jurusan_SMA = IPS	Gender	0.250
5	Rerata_SKS = range2 [18.500 - 19.500]	Gender	0.300
6	Rerata_SKS = range3 [19.500 - ∞]	Gender	0.250
7	Jurusan_SMA = IPA, Rerata_SKS = range3 [19.50...	Gender	0.100
8	Rerata_SKS = range2 [18.500 - 19.500], Jurusan_...	Gender	0.100
9	Rerata_SKS = range2 [18.500 - 19.500], Asisten	Gender	0.050
10	Rerata_SKS = range2 [18.500 - 19.500], Jurusan_...	Gender	0.050
11	Asal_Sekolah, Jurusan_SMA = IPS	Gender	0.100
12	Asal_Sekolah, Rerata_SKS = range3 [19.500 - ∞]	Gender	0.150
13	Asal_Sekolah, Jurusan_SMA = LAIN	Gender	0.050
14	Jurusan_SMA = IPS, Rerata_SKS = range3 [19.50...	Gender	0.100
15	Asisten, Rerata_SKS = range3 [19.500 - ∞]	Gender	0.150
16	Asisten, Jurusan_SMA = LAIN	Gender	0.050
17	Rerata_SKS = range3 [19.500 - ∞], Jurusan_SMA ...	Gender	0.050

Min. Criterion: confidence

Min. Criterion Value: 0.05





MODUL 12

Kegiatan

1. Mencari nilai t-hitung dan model regresi linear

1	NO_SISWA	NAMA	LAMA BELAJAR (JAM)	NILAI
2	S-101	JOKO	15	783
3	S-102	AGUS	18	877
4	S-103	SUSI	7	505
5	S-104	DYAH	9	860
6	S-105	WATI	15	968
7	S-106	IKA	17	793
8	S-107	EKO	10	752
9	S-108	YANTO	5	571
10	S-109	WAWAN	8	667
11	S-110	MAHMUD	15	723
12				

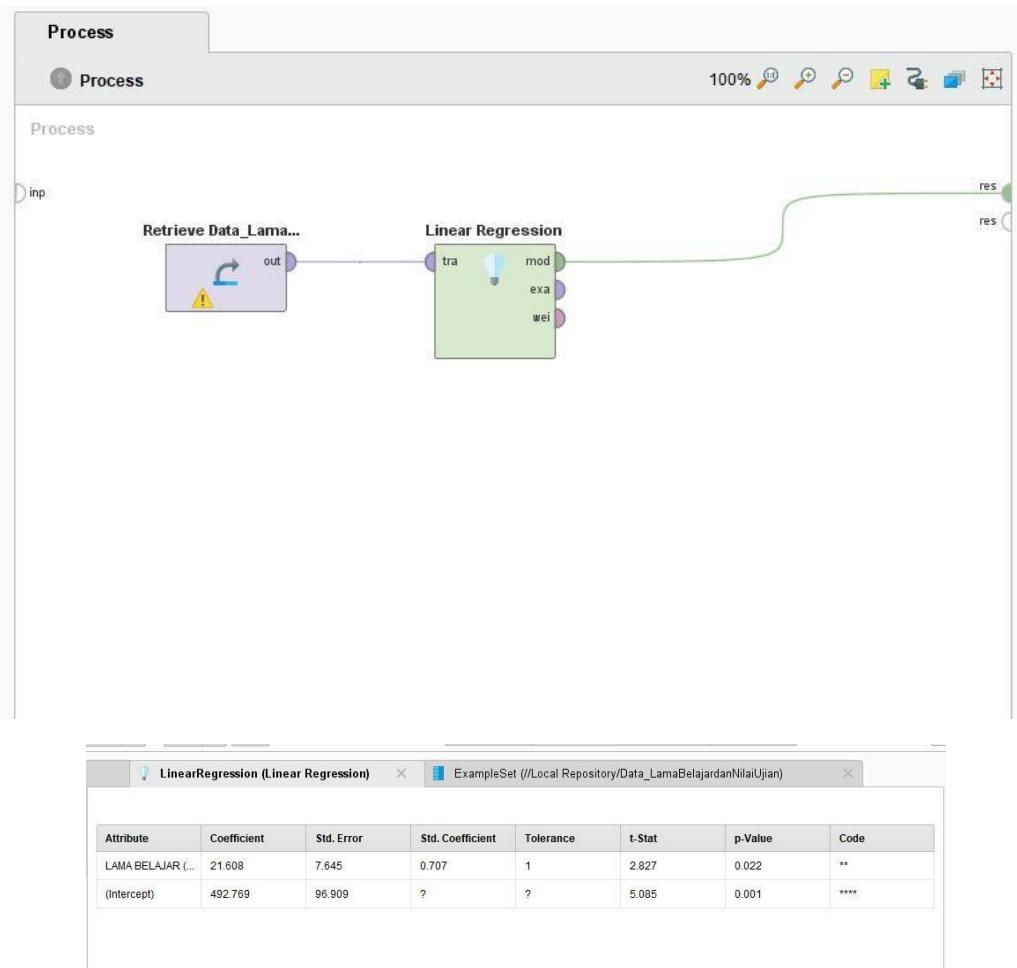
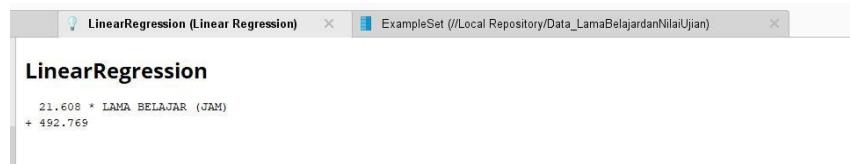


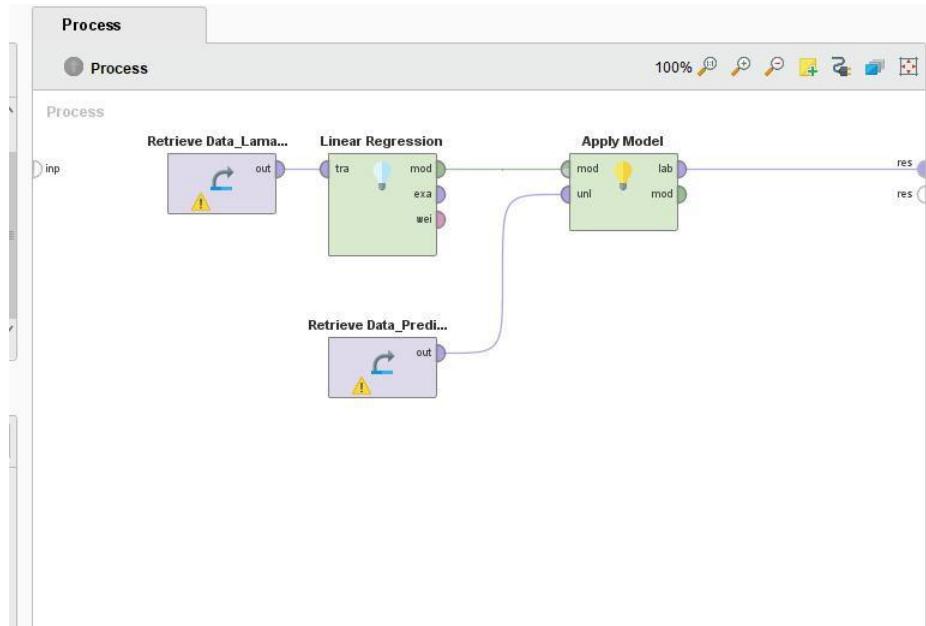
Table View

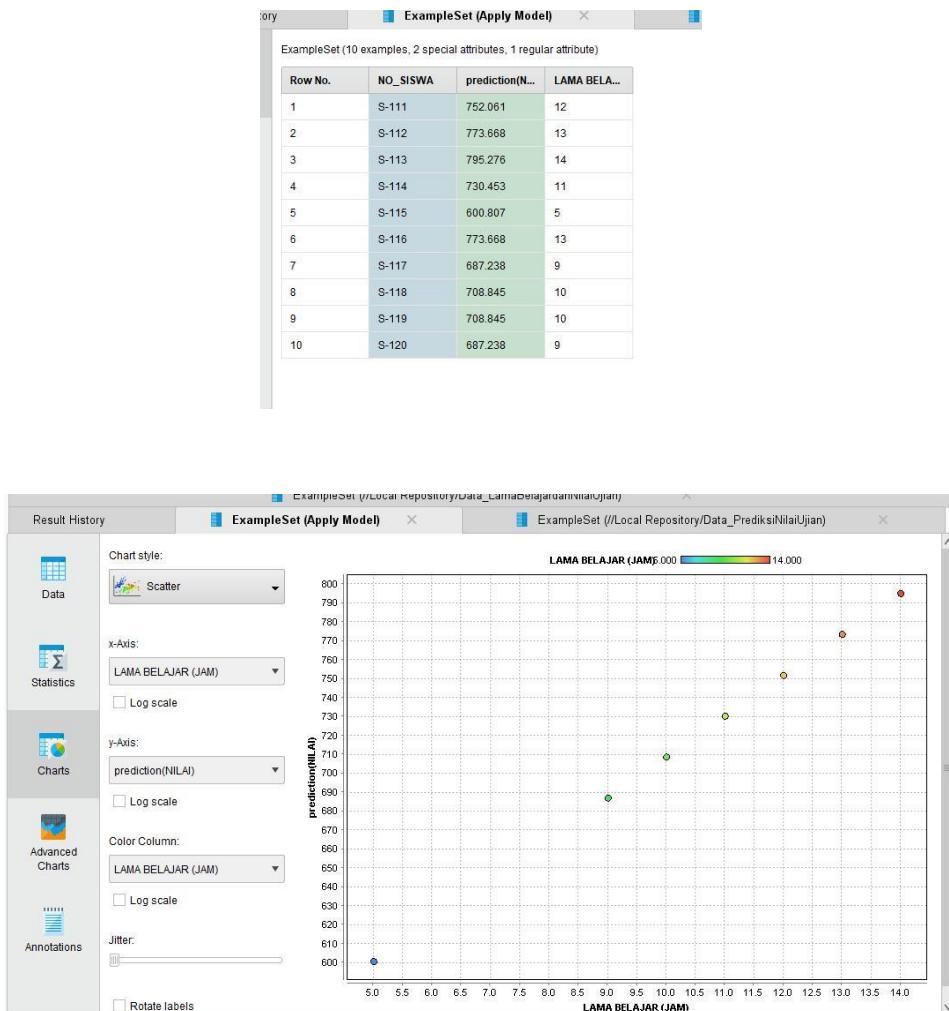


Text View

2. Mencari nilai t dan model regresi linier menggunakan RapidMiner

	A	B	C	D
1	NO_SISWA	NAMA	LAMA BELAJAR (JAM)	
2	S-111	BUDI	12	
3	S-112	SANTI	13	
4	S-113	DIAN	14	
5	S-114	DANI	11	
6	S-115	AHMAD	5	
7	S-116	BAYU	13	
8	S-117	RISA	9	
9	S-118	RANI	10	
10	S-119	YANI	10	
11	S-120	RATIH	9	
12				





3. Pembuktian Model Regresi

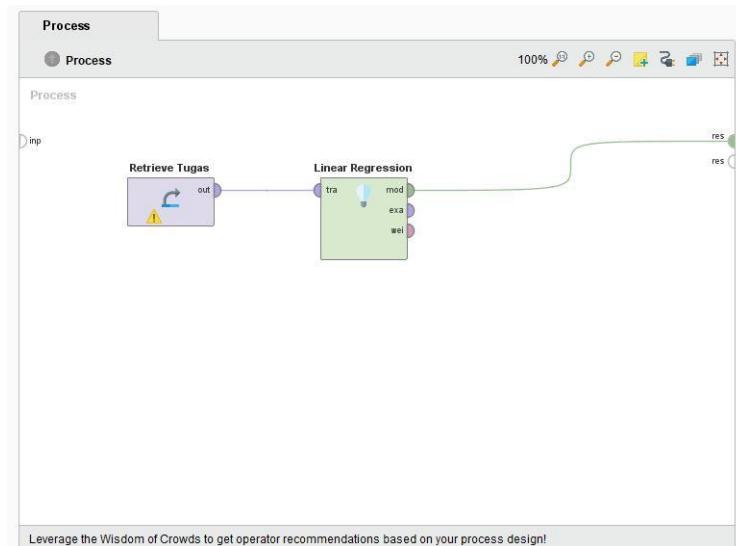
	A	B	C	D	E	F
1	NO_SISWA	NAMA	LAMA BELAJAR (JAM)	Prediction(NILAI)	Tabel	Prediction(NILAI)
2						Model Regresi
3	S-111	BUDI	12	752,061	752,065	
4	S-112	SANTI	13	773,668	773,673	
5	S-113	DIAN	14	795,276	795,281	
6	S-114	DANI	11	730,453	730,457	
7	S-115	AHMAD	5	600,807	600,809	
8	S-116	BAYU	13	773,668	773,673	
9	S-117	RISA	9	687,238	687,241	
10	S-118	RANI	10	708,845	708,849	
11	S-119	YANI	10	708,845	708,849	
12	S-120	RATIH	9	687,238	687,241	
13						

Tugas

1. Excel

A	B	C	D	E
NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)	
1	1.000.000	6	834.000	
2	1.400.000	7	1.200.000	
3	200.000	3	134.000	
4	1.400.000	6	1.167.000	
5	500.000	3	334.000	
6	1.700.000	5	1.360.000	
7	400.000	3	267.000	
9	1.900.000	5	1.520.000	
10	300.000	3	200.000	
11	500.000	4	375.000	
12	700.000	7	600.000	
13	1.900.000	3	1.267.000	
14	800.000	4	600.000	
15	1.500.000	4	1.125.000	
16	1.300.000	7	1.115.000	
17				
18				
19				

2. Proses Regresi Linier Sederhana



a. Table View

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

b. Text View

Data	LinearRegression	Description
	0.739 * PENDAPATAN (RUPIAH) + 47807.624 * JUMLAH ANGGOTA KELUARGA - 180222.487	

3. Berdasarkan aturan statistik, variable x tidak mempengaruhi secara signifikan terhadap y karena nilai t-hitung > t-table

4. Model persamaan regresi linier sederhana yang terbentuk

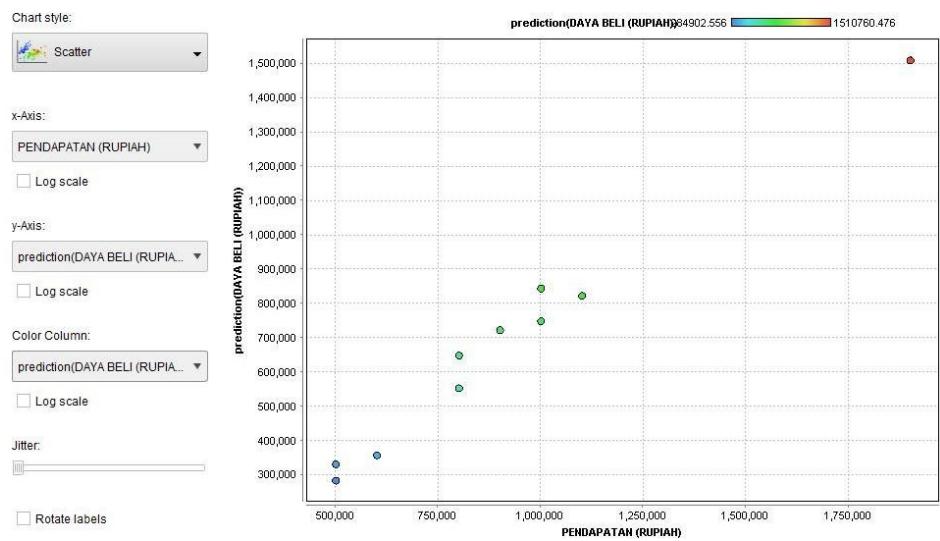
Data	LinearRegression	Description
	0.739 * PENDAPATAN (RUPIAH) + 47807.624 * JUMLAH ANGGOTA KELUARGA - 180222.487	

5. Data Testing

A	B	C	D	E	F
NO. RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	Prediction(Daya Beli(Rupiah)) Tabel	Prediction(Daya Beli(Rupiah)) Model Regresi	
1					
2					
3	1	900.000	5	723933,263	723915,633
4	2	800.000	3	554416,056	554400,385
5	3	500.000	2	284902,556	284892,761
6	4	1.900.000	6	1510760,476	1510723,257
7	5	600.000	2	358804,515	358792,761
8	6	800.000	5	650031,304	650015,633
9	7	1.000.000	6	845642,845	845623,257
10	8	1.100.000	4	823929,557	823908,009
11	9	1.000.000	4	750027,598	750008,009
12	10	500.000	3	332700,179	332700,385
13					
14					

6. Plot View

- x-Axis = Pendapatan (Rupiah),
y-Axis = Prediction (Daya Beli (Rupiah)),
Color Column = Prediction (Daya Beli (Rupiah))



- b) x-Axis = Jumlah Anggota Keluarga,
y-Axis = Prediction (Daya Beli (Rupiah)),
Color Column = Prediction (Daya Beli (Rupiah))

