

Nama : Nurul Arifia Safitri

NIM : L200170088

Kelas : D

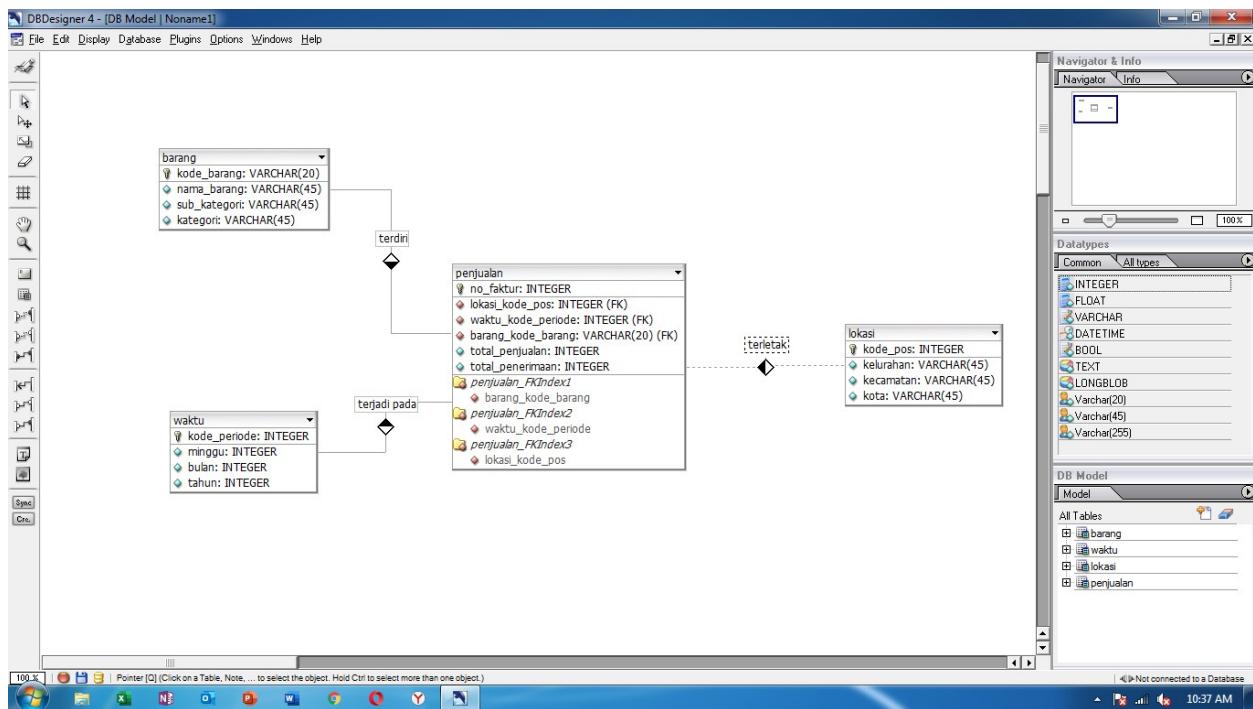
## GABUNGAN MODUL

### PRAKTIKUM DATA WERE HOUSE DAN DATA MINING

#### A. MODUL 1

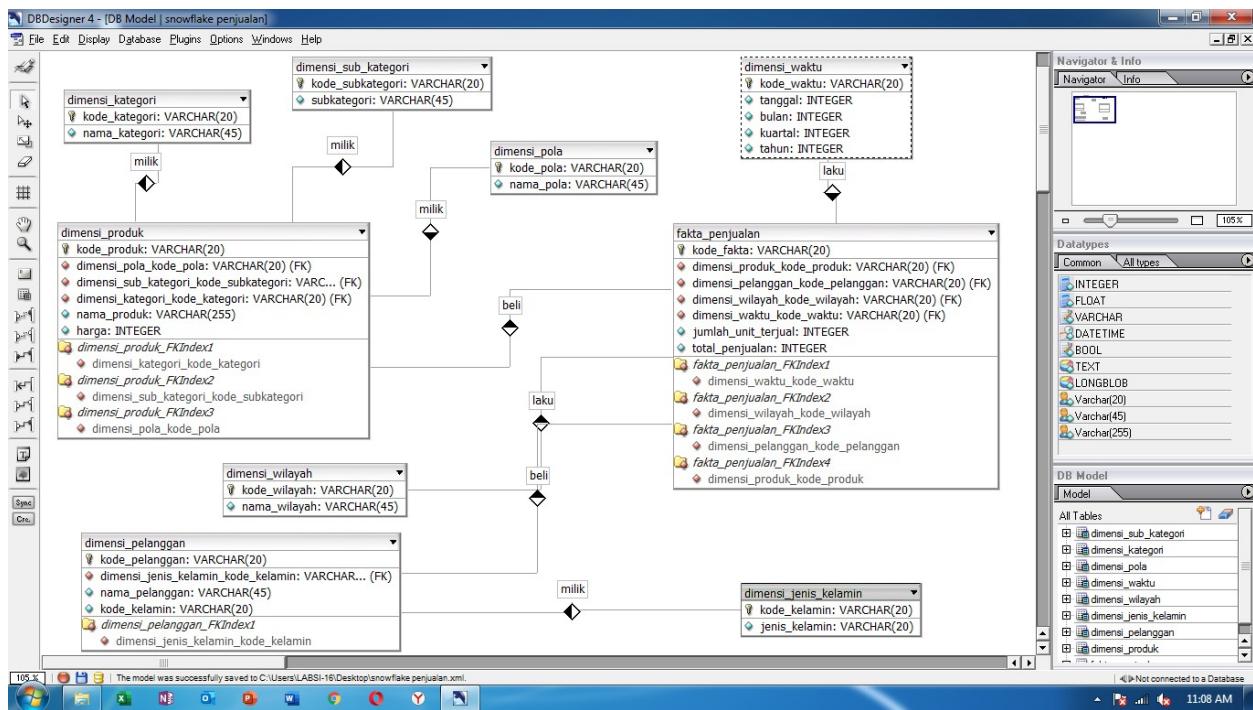
#### LATIHAN

Membuat star schema pada aplikasi DBD Desaigner dengan membuat 4 tabel barang, waktu, lokasi dan penjualan.



## TUGAS

Merancang diagram Snowflake schema dengan menggunakan DBDesigner.



## B. MODUL 5

### Kegiatan Praktikum

1. Membuat data Excel dan menyimpannya dengan nama 'Fakta\_Penjualan.xls'.



2. Membuat Pivot Table dengan menggunakan data tersebut.

The screenshot shows the 'Create PivotTable' dialog box open in Microsoft Excel. The dialog box is titled 'Create PivotTable' and contains three main sections:

- Choose the data that you want to analyze:**
  - Select a table or range
  - Use an external data source
  - Use this workbook's Data Model
- Choose where you want the PivotTable report to be placed:**
  - New Worksheet
  - Existing Worksheet
- Choose whether you want to analyze multiple tables:**
 Add this data to the Data Model

The 'Table/Range' field is set to 'Sheet1!\$A\$1:\$K\$21'. The 'Location' field is set to 'New Worksheet'. The 'OK' button is highlighted.

3. Menampilkan data Pivot Table dengan mencoba kombinasi tahun, nama\_subkategori, dan jumlah.

The screenshot shows a PivotTable named 'PivotTable1' on Sheet2. The PivotTable has 'Row Labels' set to 'tahun' and 'name\_subkategori'. The 'Sum of jumlah' value is displayed in the 'Grand Total' cell (cell A16).

The PivotTable Fields pane on the right shows the following fields:

- Choose fields to add to report:
- Search:
- Available fields:
  - bulan
  - kuartal
  - tahun
  - nama\_produk
  - nama\_subkategori
- Drag fields between areas below:
- Filters:  tahun
- Rows:  nama\_subkat...
- Values:  Sum of jumlah

## D. 2. Kegiatan 2 : Menambahkan Tipe Summary Baru

1. Menambahkan filed jumlah ke kotak Value (karena sudah ada field jumlah sebelumnya, maka field dengan nama yang sama di kotak yang sama akan memiliki tambahan karakter berupa angka di akhiran kata, pada kasus ini jumlah2 atau sum of jumlah2).

The screenshot shows a Microsoft Excel spreadsheet titled "Fakta\_Penjualan - Excel". The PivotTable Tools ribbon tab is selected. The PivotTable Fields pane on the right shows fields being mapped to the PivotTable. Under the "Values" category, "jumlah" is checked, and "Sum of jumlah" and "Count of jumlah" are also listed. The main table area contains data for items like Bahan, Batik, Bolero, Celana, Hem, Jam, Jarik, kaos, Rok, Sarimbit, and Standar, categorized by year (2010, 2011, 2012) and grouped by "jumlah". The table includes columns for "Sum of jumlah" and "Count of jumlah". The PivotTable Fields pane also shows other fields like "jenis\_kelamin", "nama\_wilayah", "harga", and "Pendapatan".

	B	C	D	E	F	G	H	I
	2010		2011		2012		Total Sum of jumlah	Total Count of jumlah2
Row Labels	Sum of jumlah	Count of jumlah2	Sum of jumlah	Count of jumlah2	Sum of jumlah	Count of jumlah2		
Bahan	1	1	8	1	8	2	17	4
Batik				1	1	1	1	1
Bolero					1		1	1
Celana	17	1			17	1	34	2
Hem	5	1	8	2	4	2	17	5
Jam					44	1	44	1
Jarik					4	1	4	1
kaos				1	14	1	15	2
Rok					1	1	1	1
Sarimbit				1	1		1	1
Standar		2		1			2	1
<b>Grand Total</b>	<b>23</b>	<b>3</b>	<b>21</b>	<b>7</b>	<b>93</b>	<b>10</b>	<b>137</b>	<b>20</b>

2. Mengubah jumlah nilai penjualan yang terjadi (sum) pada jumlah2, menjadi

jumlah transaksi yang terjadi (count).

The screenshot shows a Microsoft Excel spreadsheet titled "Fakta\_Perjualan - Excel". The PivotTable is set up with "Row Labels" as "jumlah" and columns for "2010" and "2011". The data includes items like Bahan, Batik, Bolero, Celana, Hem, Jam, Jarik, kaos, Rok, Sarimbit, Standar, and Grand Total. A "Value Field Settings" dialog box is open over the PivotTable, with "Source Name: jumlah" and "Custom Name: Count of jumlah2". Under "Summarize value field by", "Count" is selected. The "PivotTable Fields" pane on the right shows fields like "jenis\_kelamin", "nama\_wilayah", "jumlah" (which is checked), and "harga".

	2010	2011
5 Row Labels	Sum of jumlah	Sum of jumlah2
6 Bahan	1	1
7 Batik		8
8 Bolero		1
9 Celana	17	17
10 Hem	5	5
11 Jam		8
12 Jarik		
13 kaos		1
14 Rok		
15 Sarimbit		1
16 Standar		2
17 Grand Total	23	23
18		21
19		
20		
21		
22		
23		
24		
25		
26		

### D. 3. Kegiatan 3 : Calculated Field dan Calculated Item di Pivot Table

#### a) Calculated Field (Menambah field / kolom baru pada daftar field yang ada )

1. Membuat field baru dengan menggunakan tab Formulas -> Calculated Field, lalu membuat field baru dengan nama Pendapatan dengan formula, jumlah \* harga.

2. Setelah meng-klik tombol 'OK' maka akan muncul field baru dengan nama 'Pendapatan', cukup centang field tersebut dan data akan otomatis keluar.

	2010		2011		2012		
Row Labels	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	8
Batik			0			0	1
Bolero			0	1	1	255000	
Celana	17	1	935000			0	17
Hem	5	1	500000	8	2	4960000	4
Jam			0			0	44
Jarik			0			0	4
kaos			0	1	1	60000	14
Rok			0			0	1
Sarimbit			0	1	1	150000	
Standar			0	2	1	450000	
<b>Grand Total</b>	<b>23</b>	<b>3</b>	<b>15065000</b>	<b>21</b>	<b>7</b>	<b>30030000</b>	<b>93</b>

		Total Sum of jumlah	Total Count of jumlah	2	Total Sum of Pendapatan
Count of jumlah	2	Sum of Pendapatan			
	2	2120000	17	4	15045000
	1	150000	1	1	150000
		0	1	1	255000
	1	935000	34	2	3740000
	2	1596000	17	5	19023000
	1	3520000	44	1	3520000
	1	160000	4	1	160000
	1	420000	15	2	1350000
	1	225000	1	1	225000
		0	1	1	150000
		0	2	1	450000
	<b>10</b>	<b>115692000</b>	<b>137</b>	<b>20</b>	<b>456073000</b>

#### D. 4. Kegiatan 4 : Operasi Roll Up dan Drill Down

- Menghilangkan beberapa field, sehingga tinggal tersisa tahun (Columns), nama\_subkategori (Rows), dan Sum of Pendapatan (Value)

The screenshot shows a Microsoft Excel spreadsheet titled "Fakta\_Penjualan - Excel". A PivotTable is being used to analyze sales data. The PivotTable Fields pane on the right lists fields: nama\_wilayah, jumlah, harga, Pendapatan (which is checked and selected), and More Tables... The Filters pane shows "tahun" assigned to the Columns area. The main table area displays data for various products (Bahan, Batik, Bolero, Celana, Hem, Jam, Jarik, kaos, Rok, Sarimbit, Standar) across years 2010, 2011, 2012, and Grand Total. The PivotTable Tools ribbon tab is selected.

	2010	2011	2012	Grand Total
Bahan	500000	960000	2120000	15045000
Batik	0	0	150000	150000
Bolero	0	255000	0	255000
Celana	935000	0	935000	3740000
Hem	500000	4960000	1596000	19023000
Jam	0	0	3520000	3520000
Jarik	0	0	160000	160000
kaos	0	60000	420000	1350000
Rok	0	0	225000	225000
Sarimbit	0	150000	0	150000
Standar	0	450000	0	450000
<b>Grand Total</b>	<b>15065000</b>	<b>30030000</b>	<b>115692000</b>	<b>456073000</b>

- Menambahkan field pada kotak Column (kuartal dan bulan) dan Rows (nama\_produk dan nama\_pelanggan), untuk menambah field-field yang bisa

diperinci dan diringkas sesuai urutan kategori data yang lebih spesifik.

## Roll Up

PivotTable Name: Active Field: tahun

	2010	2011	2012	Grand Total
<b>Sum of Pendapatan</b>				
Bahan	500000	960000	2120000	15045000
Batik	0	0	150000	150000
Bolero	0	255000	0	255000
Celana	935000	0	935000	3740000
Hem	500000	4960000	1596000	19023000
Jam	0	352000	352000	
Jarik	0	0	160000	160000
kaos	0	60000	420000	1350000
Rok	0	0	225000	225000
Sarimbit	0	150000	0	150000
Standar	0	450000	0	450000
<b>Grand Total</b>	<b>15065000</b>	<b>30030000</b>	<b>115692000</b>	<b>456073000</b>

## Drop Down

PivotTable Name: Active Field: nama\_subkategor

	2010	2011	2012	Grand Total
<b>Sum of Pendapatan</b>				
Bahan	500000	960000	2120000	15045000
Bahan Belundru Cap Mahkota	0	0	945000	945000
Ibu Tyas	0	0	945000	945000
Bahan Standar Cap Garis	500000	0	0	500000
Ibu Tyas	500000	0	0	500000
Bahan Standar Cap Lasem	0	0	130000	130000
Ibu Niken	0	0	130000	130000
Hem Sutra Print Rama	0	960000	0	960000
Ibu Siti Arya	0	960000	0	960000
Batik	0	0	150000	150000
Celana Standar Cap Warna	0	0	150000	150000
Bapak Heru	0	0	150000	150000
Bolero	0	255000	0	255000
Kaos Batik Cap Tulis	0	255000	0	255000
Ibu Hatamah	0	255000	0	255000
Celana	935000	0	935000	3740000
Hem Standar Cap Tumpal	935000	0	0	935000
Ibu Hadi Sukarni	935000	0	0	935000
Sarimbit Standar Print Lukis	0	0	935000	935000
Ibu Hadi Sukarni	0	0	935000	935000
Hem	500000	4960000	1596000	19023000

## E. Tugas

1. Dengan menggunakan PivotTable pada file Fakta\_Penjualan.xls tambahkan 2 buah field, yaitu :
  - a. PPN (Pajak Pertambahan Nilai) sebesar 10% dari tiap pendapatan pada Pivot Table.
  - b. Total Penghasilan yang dihitung dari pendapatan dikurangi dengan PPN tersebut.

	Column Labels							
	2010			2011				
Row Labels	Sum of Pendapatan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal	Sum of Pendapatan	Sum of PPN (10%)	
Bahan	500000	50000	450000	1	2	960000	96000	
Batik	0	0	0			0	0	
Bolero	0	0	0			255000	25500	
Celana	935000	93500	841500	17	1	0	0	
Hem	500000	50000	450000	5	4	4960000	496000	
Jam	0	0	0			0	0	
Jarik	0	0	0			0	0	
kaos	0	0	0			60000	6000	
Rok	0	0	0			0	0	
Sarimbit	0	0	0			150000	15000	
Standar	0	0	0			450000	45000	
<b>Grand Total</b>	<b>15065000</b>	<b>1506500</b>	<b>13558500</b>	<b>23</b>	<b>7</b>	<b>30030000</b>	<b>3003000</b>	

	2012							
	Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal	Sum of Pendapatan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of jumlah	Sum of kuartal
	864000	8	1	2120000	212000	1908000	8	4
	0			150000	15000	135000	1	3
	229500	1	4	0	0	0		
	0			935000	93500	841500	17	1
	4464000	8	5	1596000	159600	1436400	4	5
	0			3520000	352000	3168000	44	1
	0			160000	16000	144000	4	2
	54000	1	1	420000	42000	378000	14	1
	0			225000	22500	202500	1	4
	135000	1	4	0	0	0		
	405000	2	4	0	0	0		
	<b>27027000</b>	<b>21</b>	<b>19</b>	<b>115692000</b>	<b>11569200</b>	<b>104122800</b>	<b>93</b>	<b>21</b>

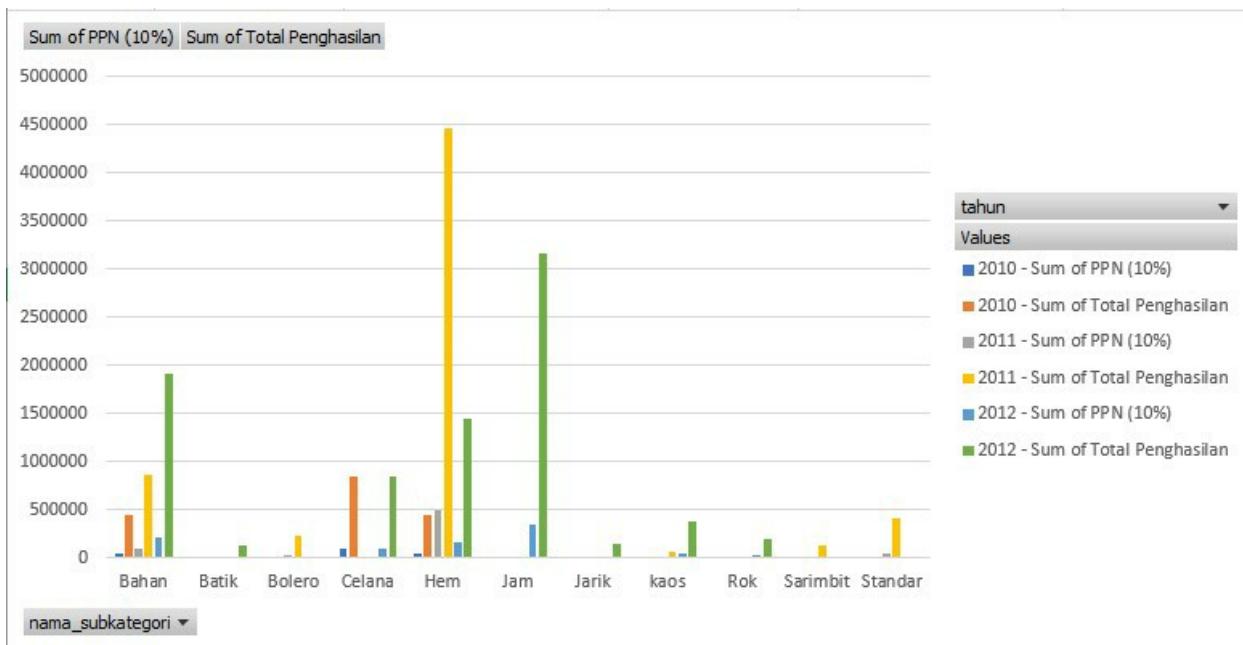
Total Sum of Pendapatan	Total Sum of PPN (10%)	Total Sum of Total Penghasilan	Total Sum of jumlah	Total Sum of kuartal
15045000	1504500	13540500	17	7
150000	15000	135000	1	3
255000	25500	229500	1	4
3740000	374000	3366000	34	2
19023000	1902300	17120700	17	14
3520000	352000	3168000	44	1
160000	16000	144000	4	2
1350000	135000	1215000	15	2
225000	22500	202500	1	4
150000	15000	135000	1	4
450000	45000	405000	2	4
<b>45607300</b>	<b>45607300</b>	<b>410465700</b>	<b>137</b>	<b>47</b>

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

PivotTable

Row Labels	Column Labels		2010		2011		2012		Total Sum of PPN (10%)	Total Sum of Total Penghasilan
	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan	Sum of PPN (10%)	Sum of Total Penghasilan		
Bahan	50000	450000	96000	864000	212000	1908000	1504500	13540500		
Batik	0	0	0	0	15000	135000	15000	135000		
Bolero	0	0	25500	229500	0	0	25500	229500		
Celana	93500	841500	0	0	93500	841500	374000	3366000		
Hem	50000	450000	496000	4464000	159600	1436400	1902300	17120700		
Jam	0	0	0	0	352000	3168000	352000	3168000		
Jarik	0	0	0	0	16000	144000	16000	144000		
kaos	0	0	6000	54000	42000	378000	135000	1215000		
Rok	0	0	0	0	22500	202500	22500	202500		
Sarimbit	0	0	15000	135000	0	0	15000	135000		
Standar	0	0	45000	405000	0	0	45000	405000		
<b>Grand Total</b>	<b>1506500</b>	<b>13558500</b>	<b>3003000</b>	<b>27027000</b>	<b>11569200</b>	<b>104122800</b>	<b>45607300</b>	<b>410465700</b>		

## Pivot chart



Dapat dilihat dari data di atas bahwa, kategori produk dengan nilai penghasilan terbanyak pada tahun 2010 – 2012 adalah Hem.

## C. MODUL 6

### TUGAS

1. Jumlah data masing-masing kelas IPA, IPS dan Lain.

25	
26	
27	10
28	6
29	4
30	LAIN

Rumus :

A screenshot of a Microsoft Excel spreadsheet. The ribbon menu is visible at the top. The formula bar shows the formula `=COUNTIF(A2:A21;"IPA")`. The cell B27 contains the value 10, which corresponds to the count of "IPA" in the range A2:A21. The rest of the cells in the row are empty.

31		
32	13	TEPAT
33	7	TERLAMBAT
34		

Rumus :

The screenshot shows a Microsoft Excel spreadsheet with the following data:

25		
26		
27	10	IPA
28	6	IPS
29	4	LAIN
30		
31		
32	13	TEPAT
33	7	TERLAMBAT
34		

The formula bar at the top shows the formula `=COUNTIF(F2:F21;"TEPAT")`. A red box highlights the formula in the formula bar.

### 3. Nilai MAX, MIN, MEAN, STANDART DEVIATION

23	MAX
16	MIN
18,95	MEAN
1,66938375	STANDAR DEVIASI

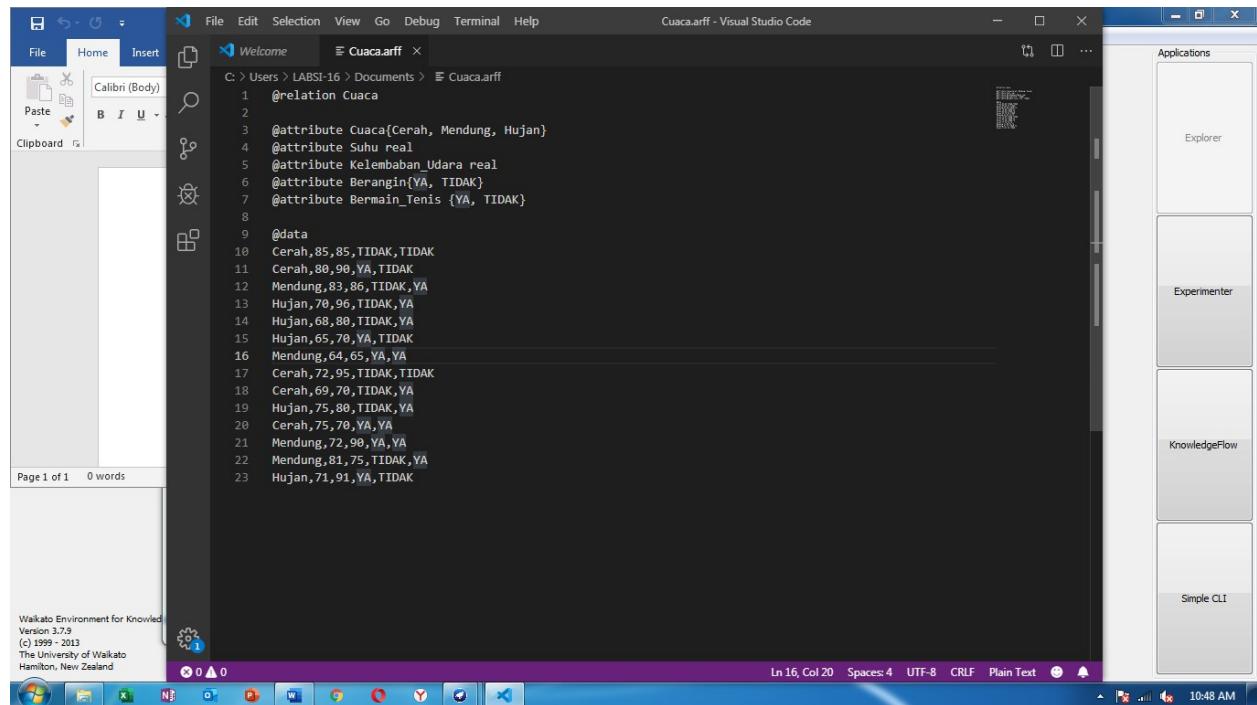
	<b>3 JUMLAH DATA GABUNGAN</b>	

Rumus :

Font	Alignment	Number
<input type="button" value="fx"/>	=COUNTIFS(A2:A21;"IPA";B2:B21;"PRIA";E2:E21;"YA";F2:F21;"TEPAT")	
B	C	D

## D. MODUL 7

### KEGIATAN PRAKTIKUM



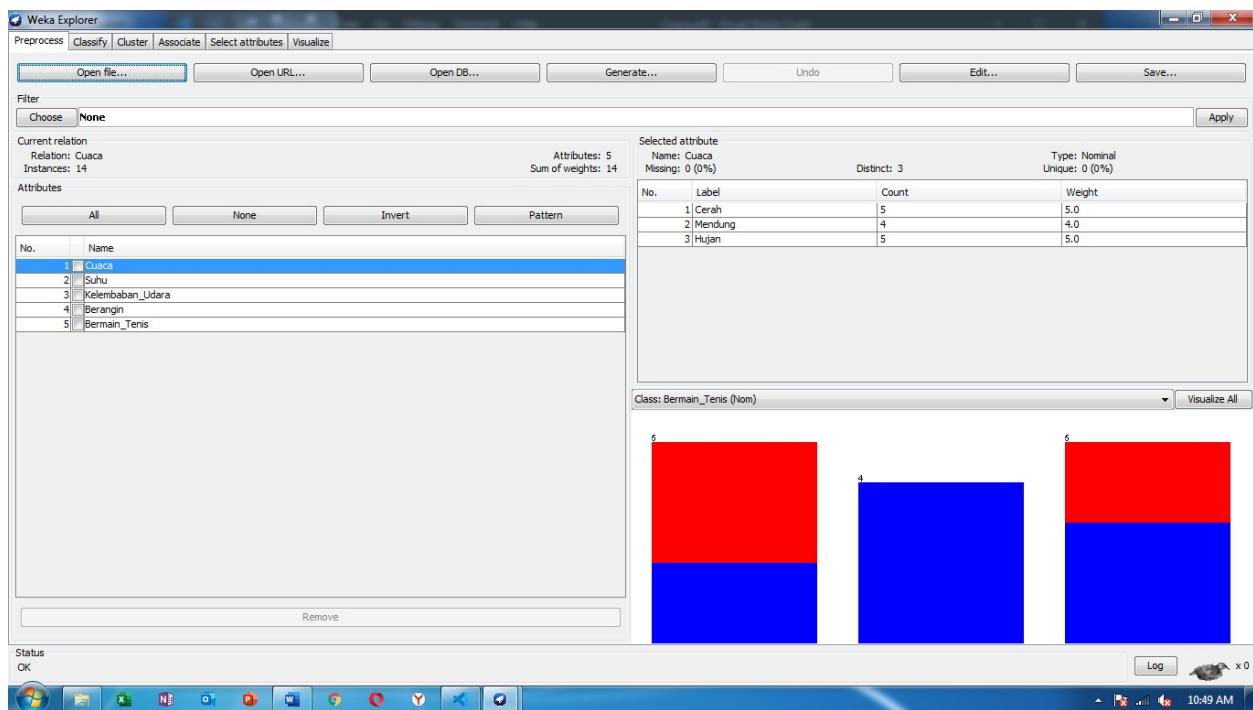
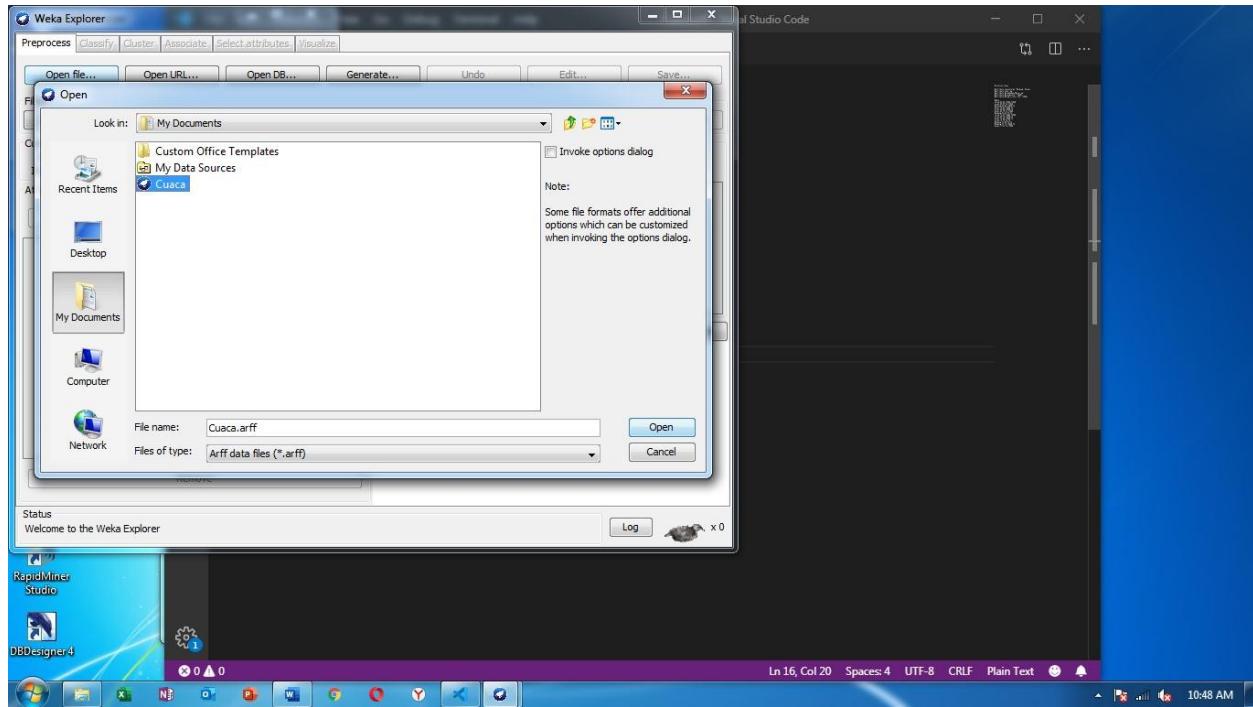
```

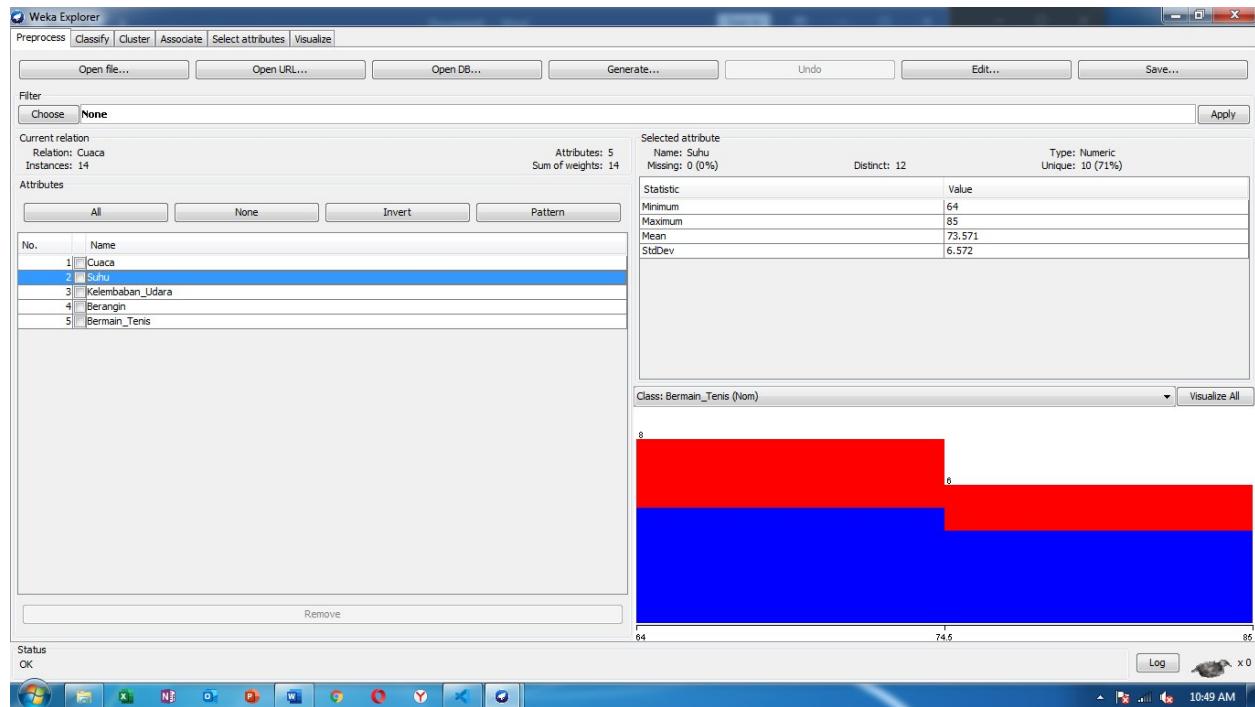
@relation Cuaca
@attribute Suhu real
@attribute Kelembaban_Udara real
@attribute Berangin{YA, TIDAK}
@attribute Bermani_Tenis {YA, TIDAK}

@data
Cerah,85,85,TIDAK,TIDAK
Cerah,80,90,YA,TIDAK
Mendung,83,86,TIDAK,YA
Hujan,78,96,TIDAK,YA
Hujan,68,80,TIDAK,YA
Hujan,65,70,YA,TIDAK
Mendung,64,65,YA,YA
Cerah,72,95,TIDAK,TIDAK
Cerah,69,70,TIDAK,YA
Hujan,75,88,TIDAK,YA
Cerah,75,70,YA,YA
Mendung,72,90,YA,YA
Mendung,81,75,TIDAK,YA
Hujan,/1,91,YA,TIDAK

```

WalkMe Environment for Knowledge  
Version 3.7.9  
(c) 1999 - 2013  
The University of Waikato  
Hamilton, New Zealand





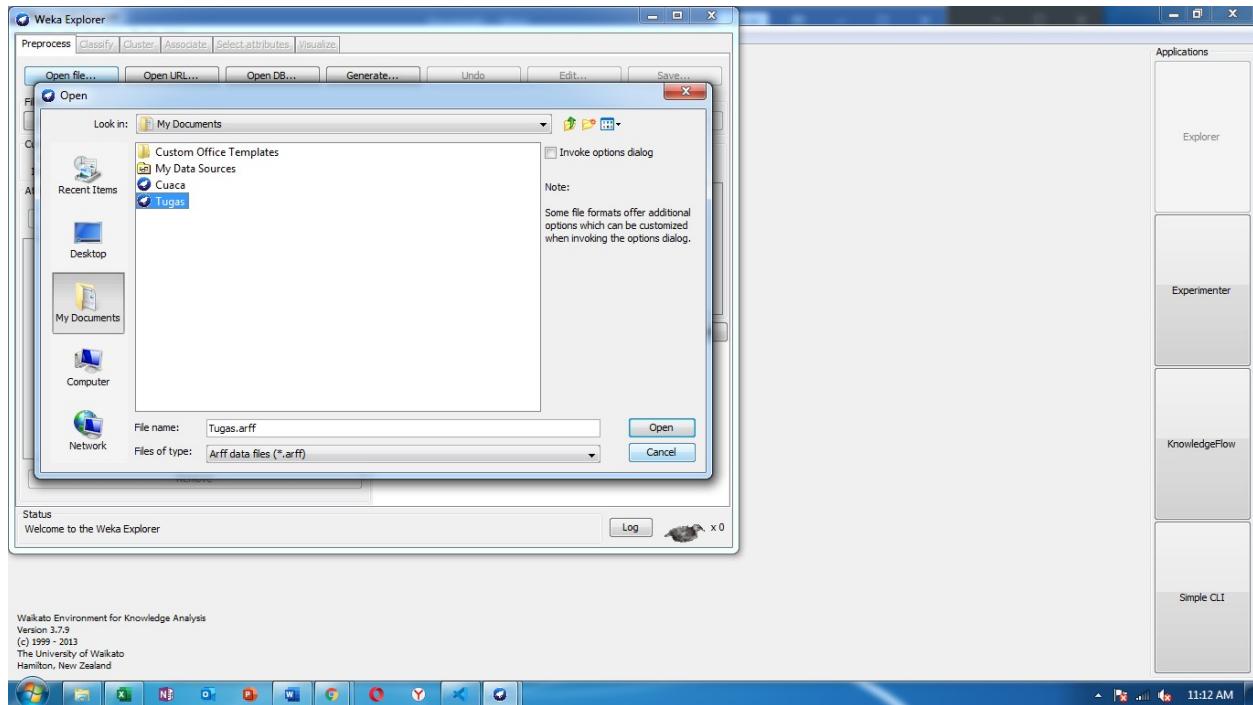
## TUGAS

1.

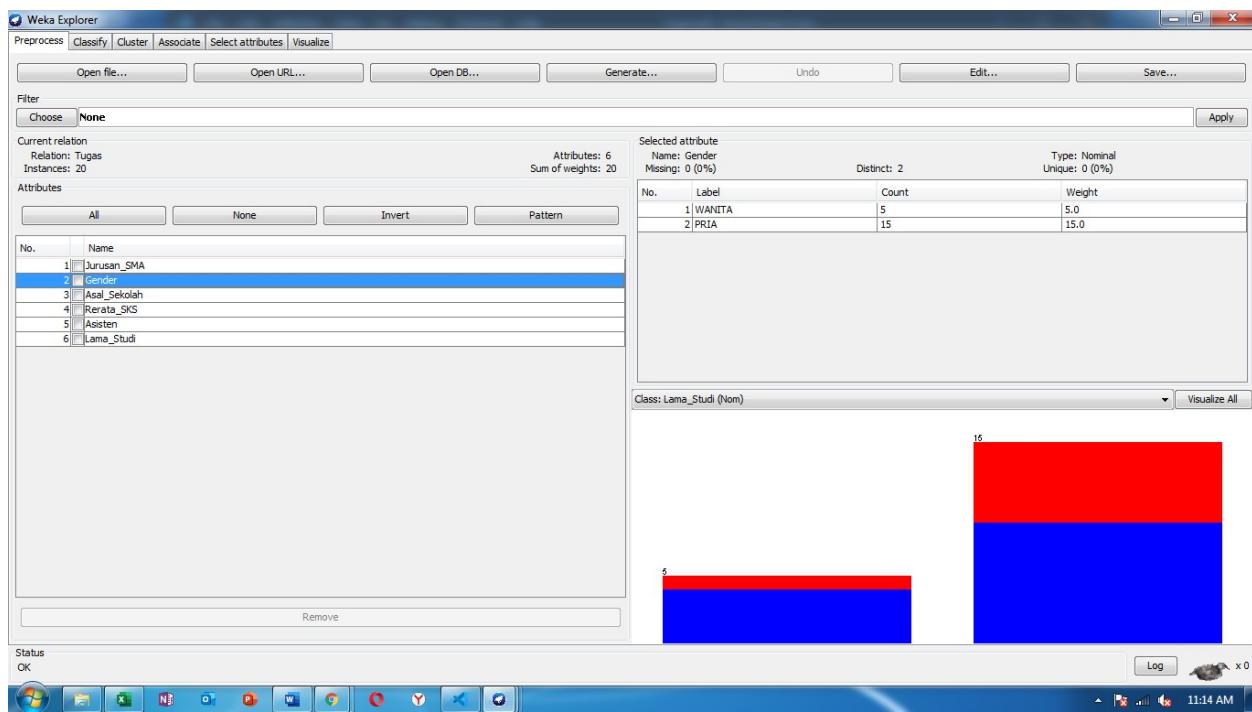
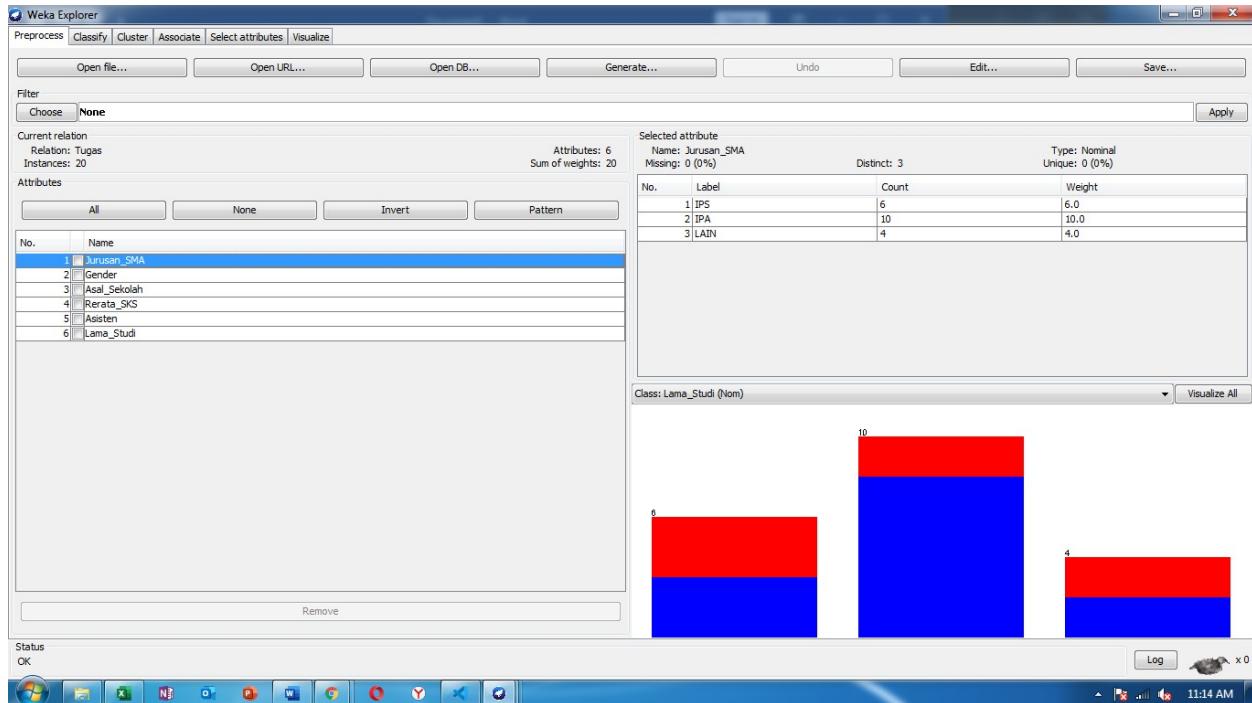
The screenshot shows the Weka Explorer interface. On the left, the 'Weka Explorer' window displays the 'Preprocess' tab selected. It shows a list of attributes: No., Name, and instances. The first instance is highlighted. The status bar at the bottom indicates 'OK'. On the right, the 'Tugas.arff - Visual Studio Code' window shows the ARFF file content. The code includes attribute definitions and a data section with 30 entries. The status bar at the bottom of this window shows 'Ln 20, Col 37' and other file details.

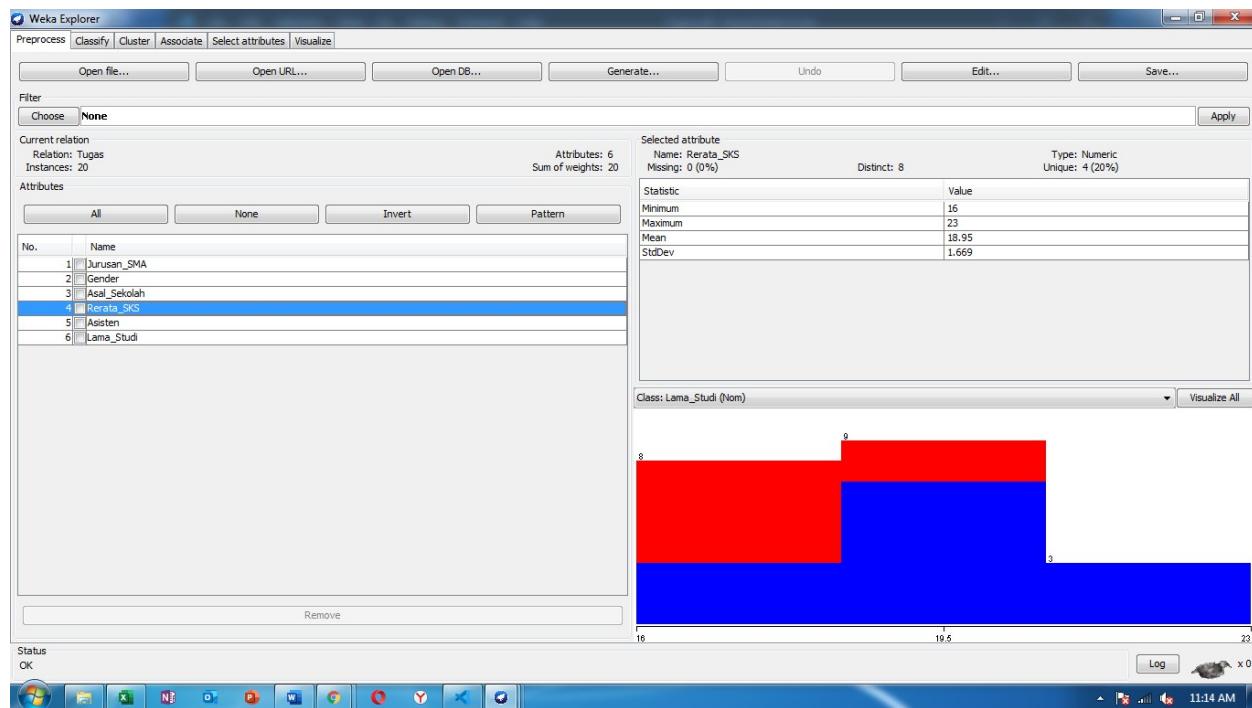
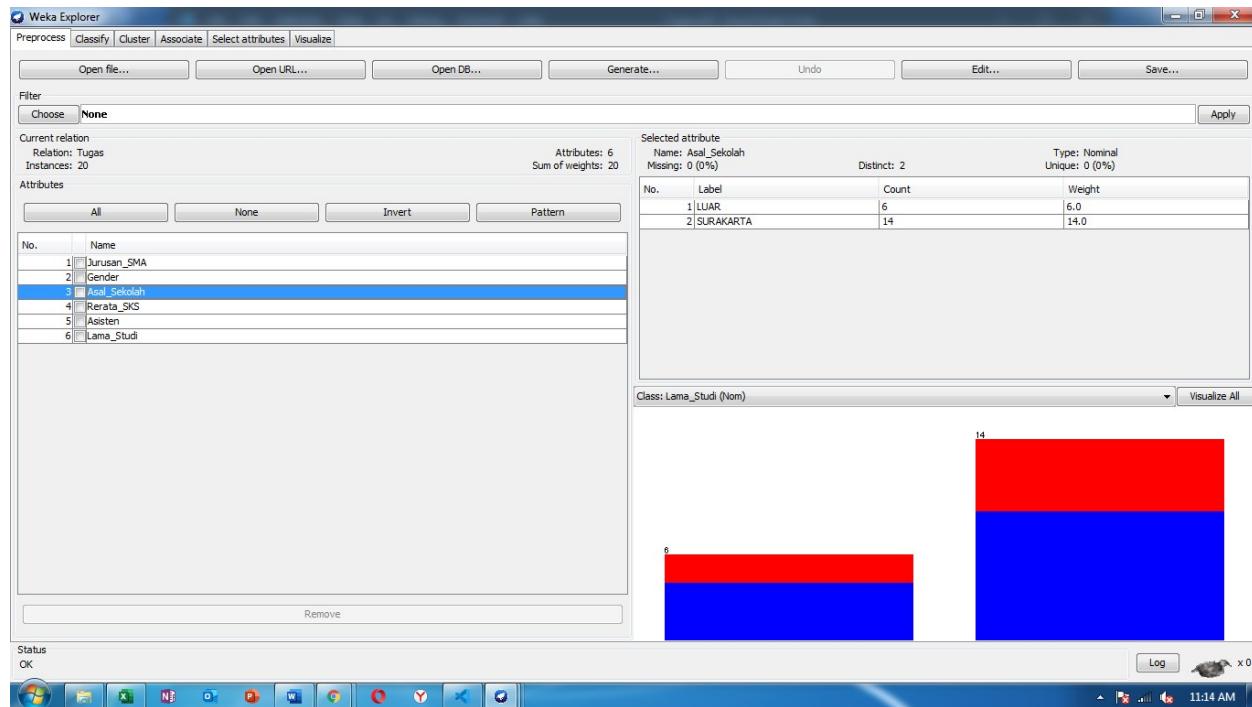
```
C:\> Users > LABSI-16 > Documents > Tugas.arff
@relation Tugas
@attribute Jurusan_SMA{IPS, IPA, LAIN}
@attribute Gender{WANITA, PRIA}
@attribute Asal_Sekolah{LUAR, SURAKARTA}
@attribute Rerata_SKS real
@attribute Asisten{TIDAK, YA}
@attribute Lama_Studi{TEPAT, TERLAMBAT}

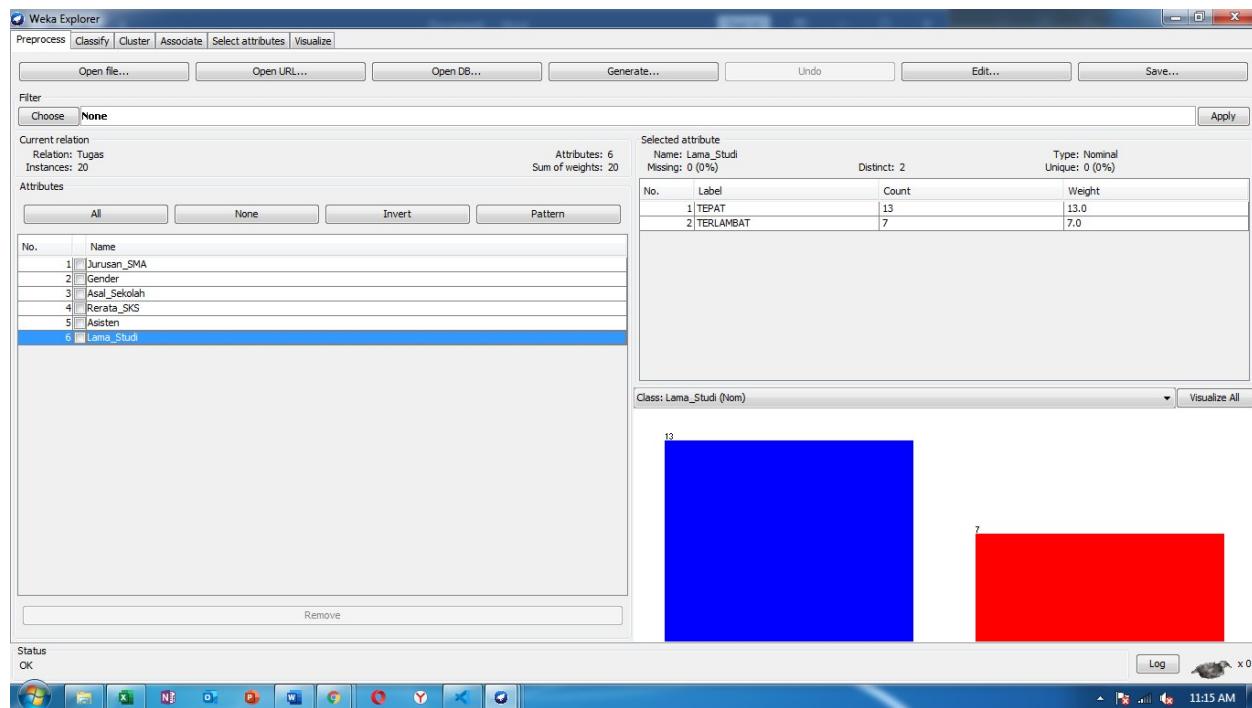
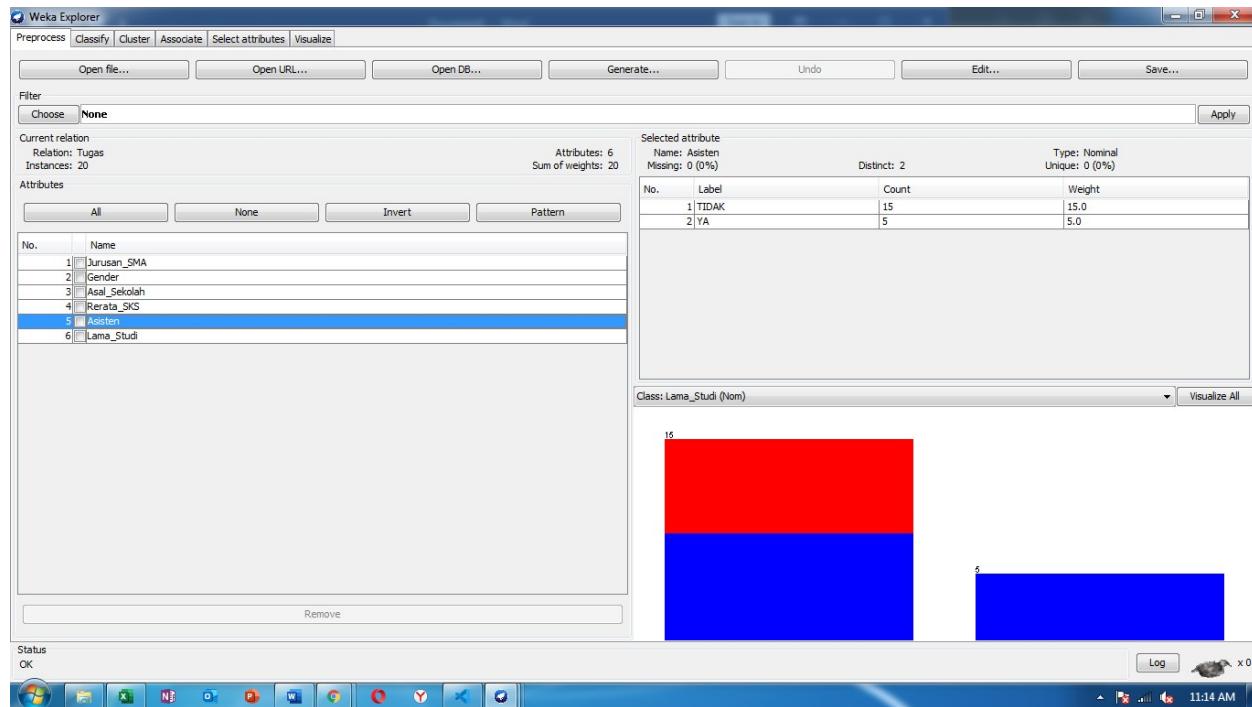
@data
IPS,WANITA,SURAKARTA,18,TIDAK,TERLAMBAT
IPA,PRIA,SURAKARTA,19,YA,TEPAT
LAIN,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
IPA,PRIA,LUAR,17,TIDAK,TERLAMBAT
IPA,WANITA,SURAKARTA,17,TIDAK,TEPAT
IPA,WANITA,LUAR,18,YA,TEPAT
IPA,PRIA,SURAKARTA,18,TIDAK,TERLAMBAT
IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPS,PRIA,LUAR,18,TIDAK,TERLAMBAT
LAIN,WANITA,SURAKARTA,18,TIDAK,TEPAT
IPA,WANITA,SURAKARTA,19,TIDAK,TEPAT
IPS,PRIA,SURAKARTA,20,TIDAK,TEPAT
IPS,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPA,PRIA,SURAKARTA,19,TIDAK,TEPAT
IPA,PRIA,LUAR,22,YA,TEPAT
LAIN,PRIA,SURAKARTA,16,TIDAK,TERLAMBAT
IPS,PRIA,LUAR,20,TIDAK,TEPAT
LAIN,PRIA,LUAR,23,YA,TEPAT
IPA,PRIA,SURAKARTA,21,YA,TEPAT
IPS,PRIA,SURAKARTA,19,TIDAK,TERLAMBAT
```



2.







3. a. Jumlah atribut bertipe binomial terdapat 4 data (Gender, Asal\_Sekolah, Asisten, Lama\_Studi)

b. Jumlah atribut bertipe polynomial terdapat 1 data (Jurusan\_SMA)

c. Jumlah atribut bertipe real terdapat 1 data (Rerata\_SKS)

4.

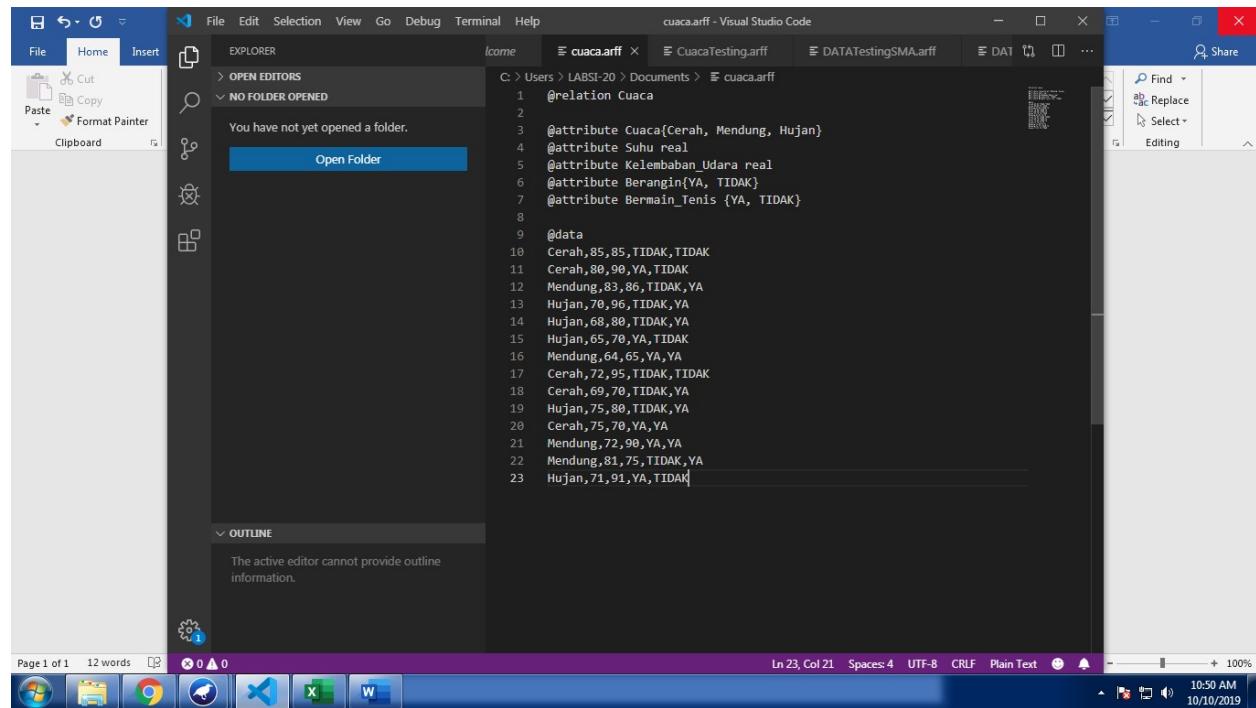
23	NILAI MAX
16	NILAI
18.95	NILAI MEAN
1.669384	STANDARD

## E. MODUL 8

### KEGIATAN PRAKTIKUM

1.

- Cuaca.arff



```
File Edit Selection View Go Debug Terminal Help cuaca.arff - Visual Studio Code
File Home Insert
EXPLORER lorne cuaca.arff CuacaTesting.arff DAT TestingSMA.arff ...
C:\> Users > LABSI-20 > 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,88,98,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,72,99,YA,YA
22 Mendung,81,75,TIDAK,YA
23 Hujan,71,91,YA,TIDAK
```

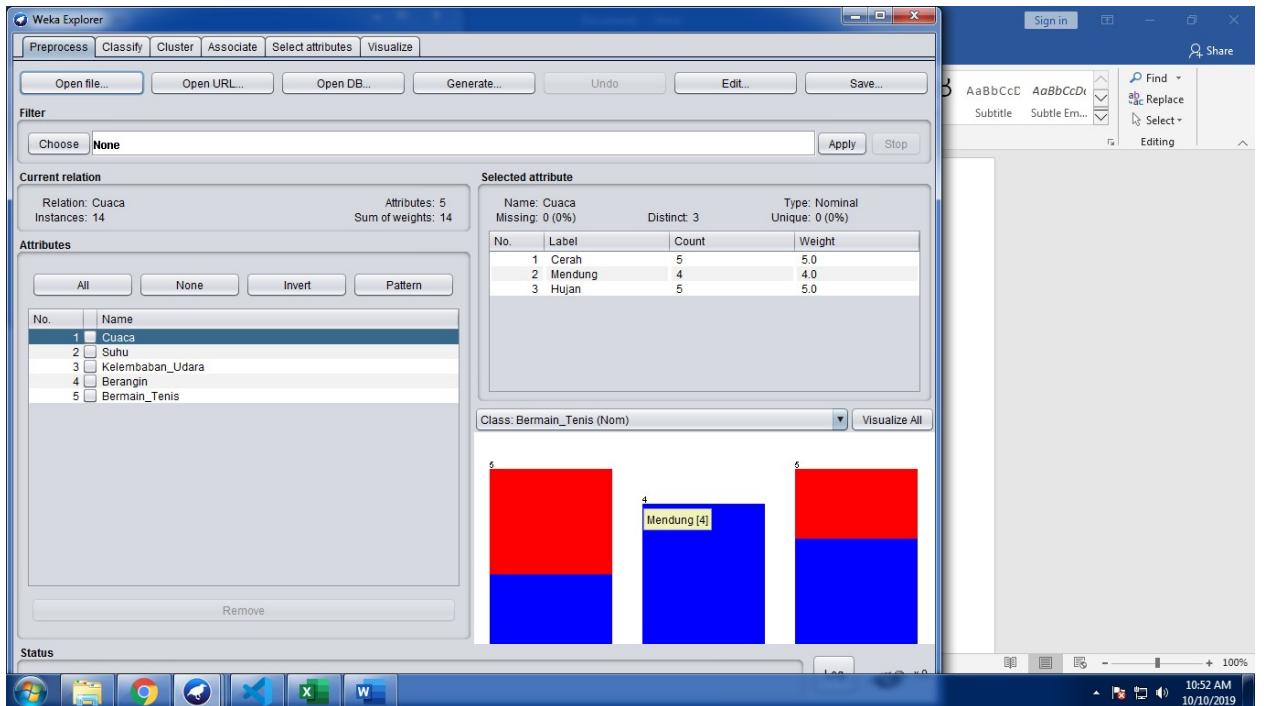
- CuacaTesting.arff

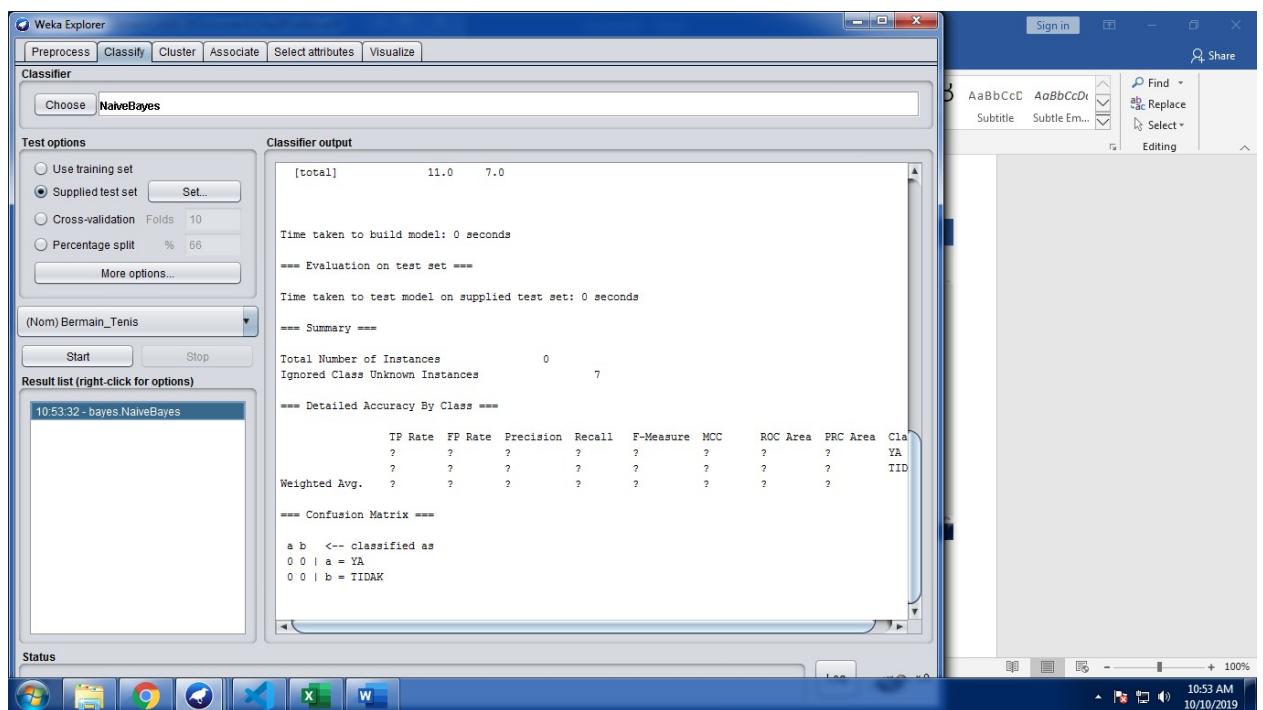
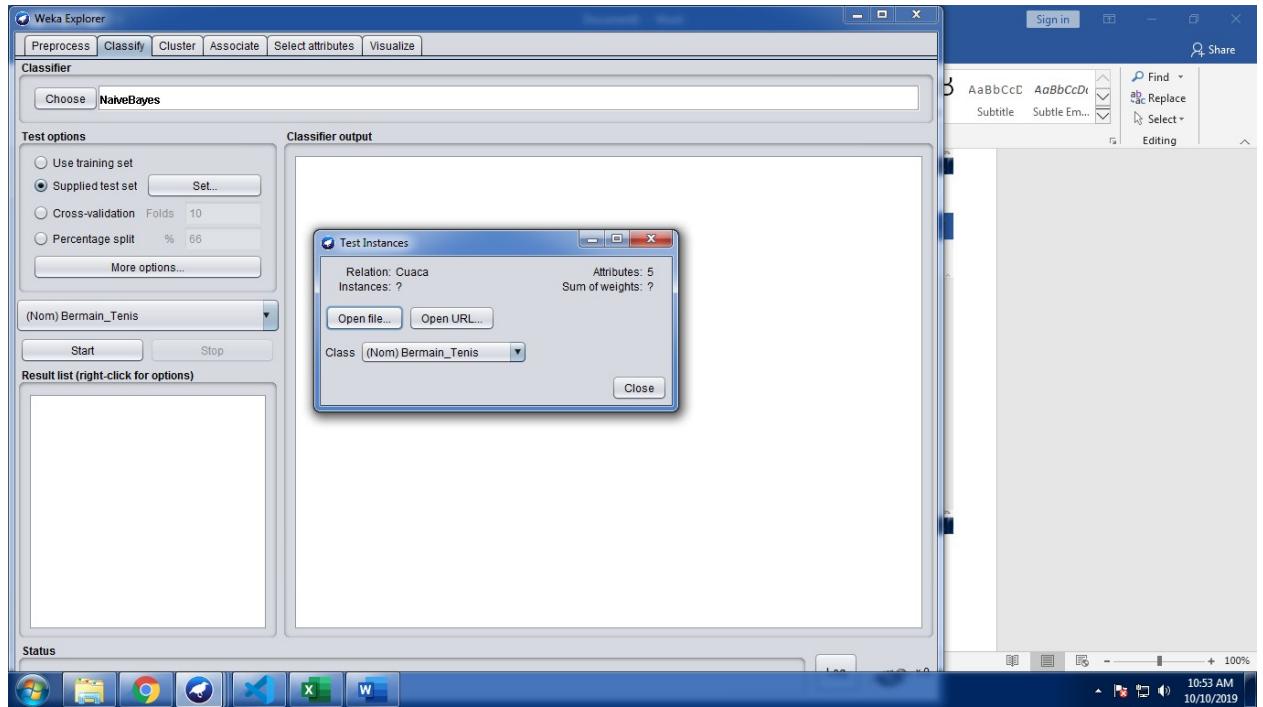
```

File Edit Selection View Go Debug Terminal Help CuacaTesting.arff - Visual Studio Code
PROTECTED VIEW Be c
EXPLORER lcome cuaca.arff CuacaTesting.arff (deleted) DATATestingSMA.arff
OPEN EDITORS NO FOLDER OPENED You have not yet opened a folder.
Open Folder
OUTLINE The active editor cannot provide outline information.
@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, ?

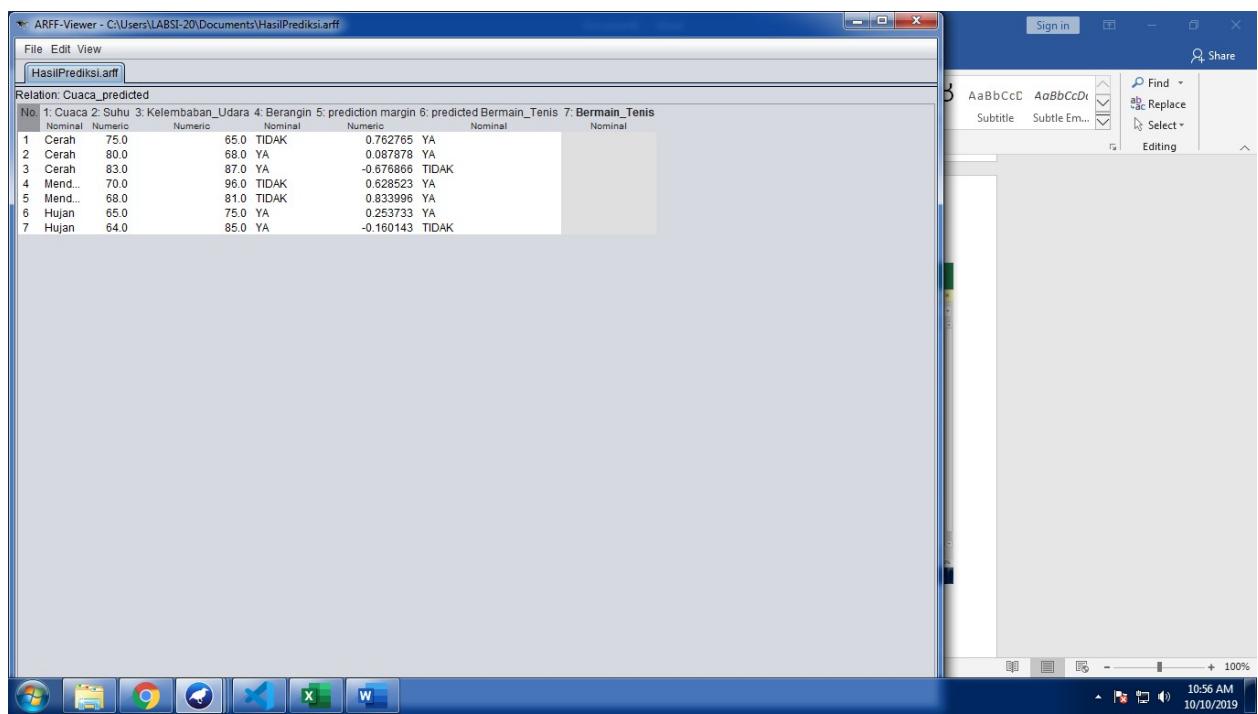
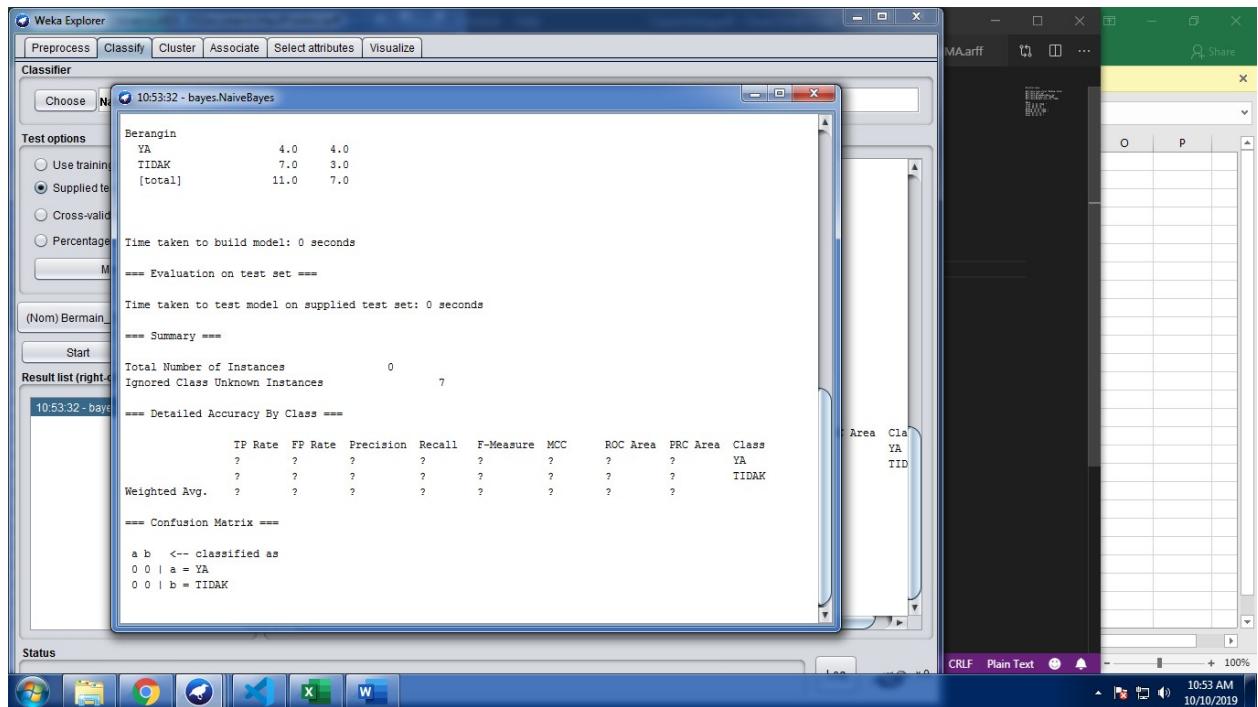
```

2.





Klik kanan lalu pilih view in separate window



PROTECTED VIEW Be careful—files from the Internet can contain viruses. Unless you need to edit, it's safer to stay in Protected View. [Enable Editing](#)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Cuaca	Suhu	Kelembaban_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										
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															
26															

Count: 5

Training Testing

10:58 AM 10/10/2019

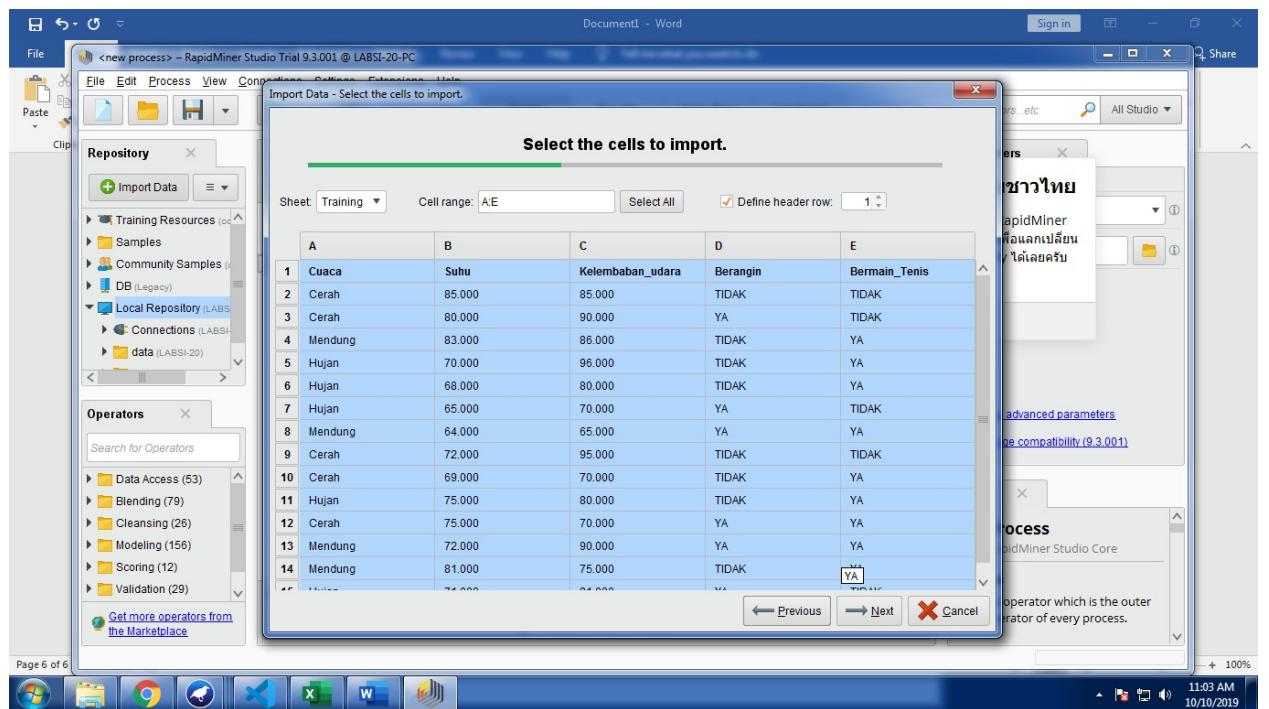
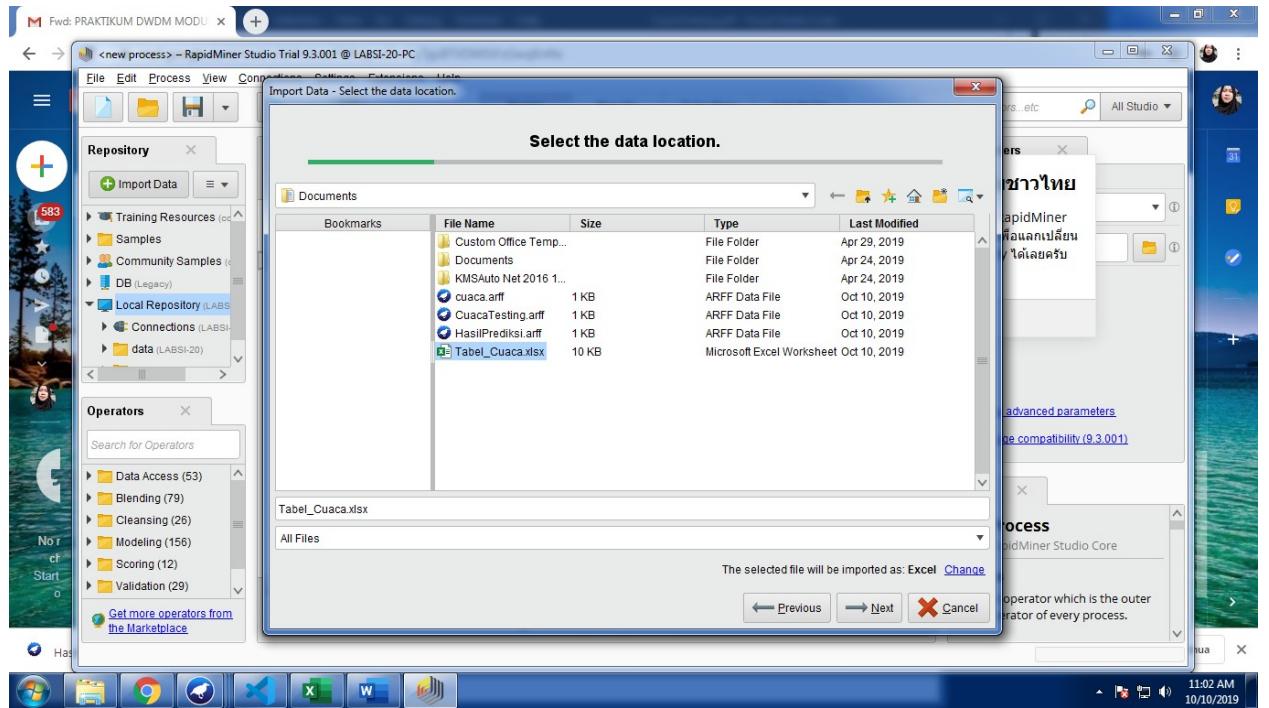
PROTECTED VIEW Be careful—files from the Internet can contain viruses. Unless you need to edit, it's safer to stay in Protected View. [Enable Editing](#)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Cuaca	Suhu	Kelembaban_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													
9																	
10																	
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	

Training Testing

10:58 AM 10/10/2019

## Data Training



Fwd: PRAKTIKUM DWDM MODU

<new process> - RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC

File Edit Process View Connections Settings Extensions Help

Repository Import Data

Training Resources Samples Community Samples DB (Legacy) Local Repository (LABSI-20) data (LABSI-20)

Operators Search for Operators Data Access (53) Blending (79) Cleansing (26) Modeling (156) Scoring (12) Validation (29) Get more operators from the Marketplace

Import Data - Format your columns.

Format your columns.

Replace errors with missing values

Cuaca polynominal Suhu integer Kelembaban\_u... integer Berangin polynominal Bermain\_Tenis polynominal

	Cuaca	Suhu	Kelembaban_u...	Berangin	Bermain_Tenis
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

Change Type > polynomial binomial real integer date\_time date advanced

Change Role Rename column Exclude column

no problems.

Previous Next Cancel

11:03 AM 10/10/2019

Document1 - Word

File Edit Process View Connections Settings Extensions Help

Clipboard Paste

Results History ExampleSet (/Local Repository/DataCuaca\_Training)

Turbo Prep Auto Model

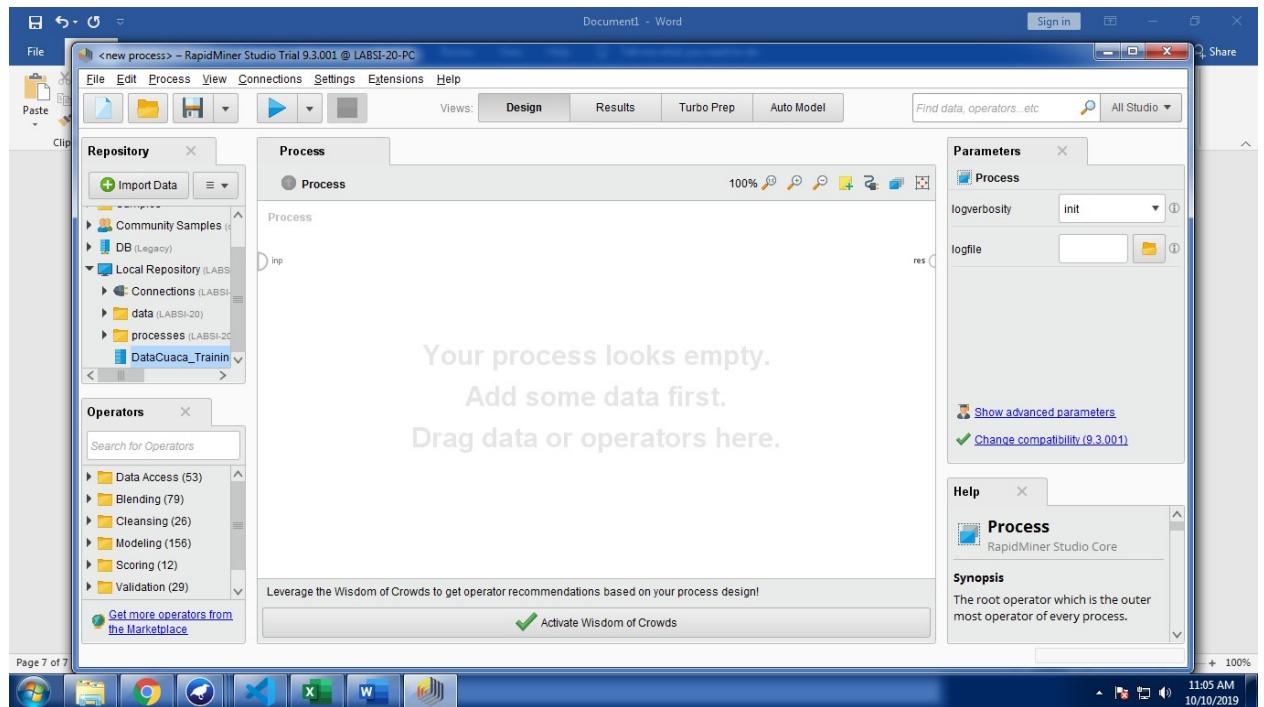
Filter (14 / 14 examples): all

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

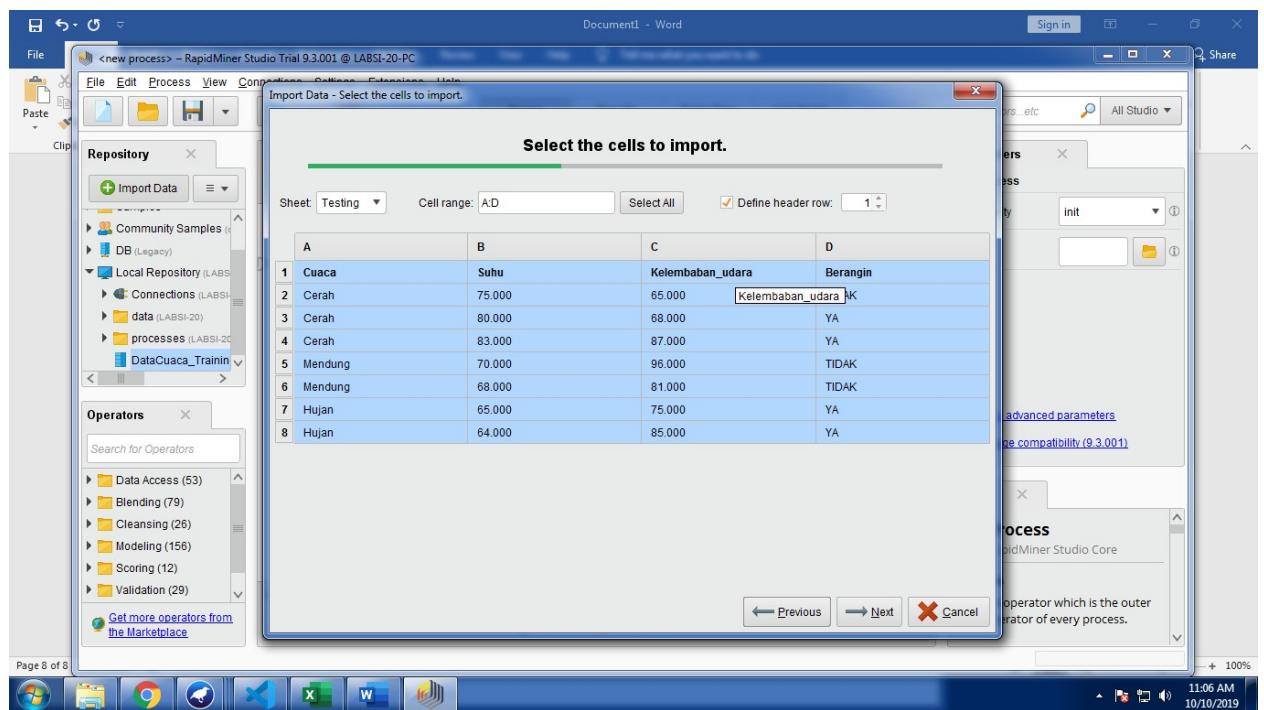
ExampleSet (14 examples, 1 special attribute, 4 regular attributes)

Page 7 of 7

11:05 AM 10/10/2019



## Data Testing



**Import Data - Format your columns.**

**Format your columns.**

Replace errors with missing values

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

no problems.

Previous Next Cancel

**Document1 - Word**

**File Edit Process View Connections Settings Extensions Help**

**Process**

Process

Process

Retrieve DataCuaca\_Training

Retrieve DataCuaca\_Testing

Parameters

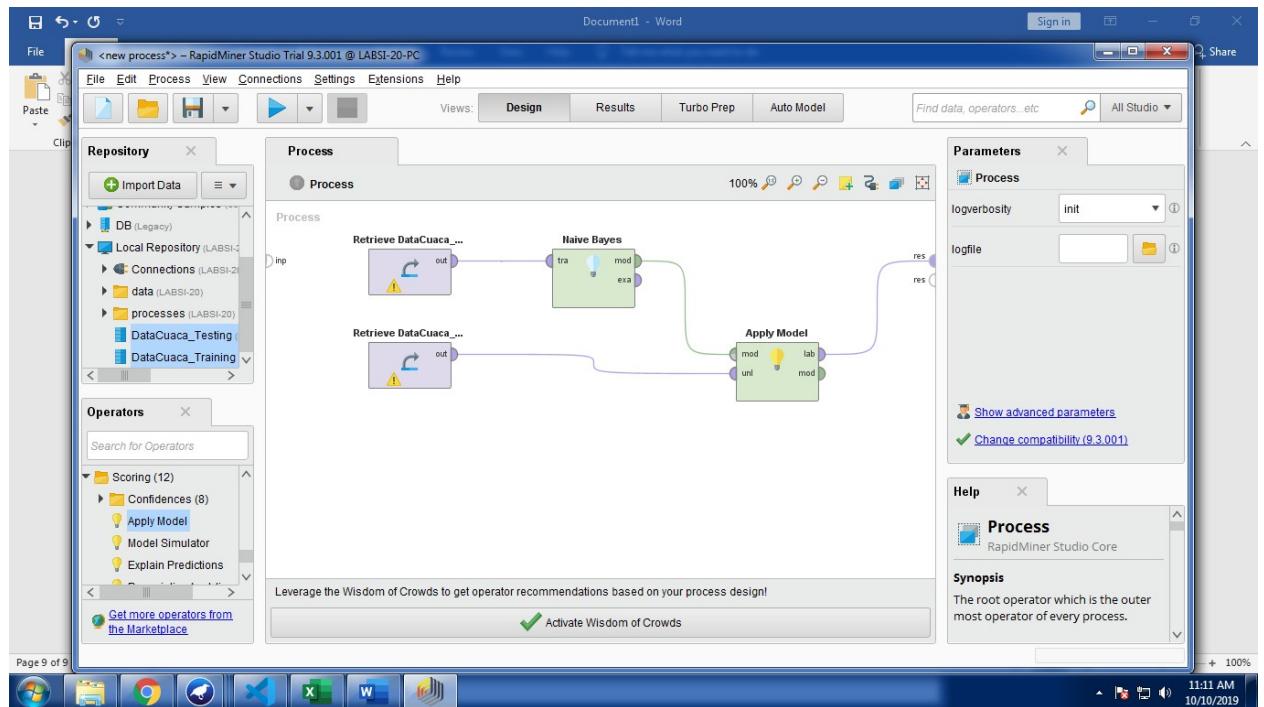
repository entry: DataCuaca\_Training

Help

Retrieval

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds



Document1 - Word

RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC

**File Edit Process View Connections Settings Extensions Help**

**Views:** Design Results Turbo Prep Auto Model

**Find data, operators... etc.** **All Studio**

**Result History**

ExampleSet (/Local Repository/DataCuaca\_Training)

ExampleSet (Apply Model)

ExampleSet (/Local Repository/DataCuaca\_Testing)

**Repository**

- Import Data
- Training Resources (connected)
- Samples
- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-20)
  - Connections (LABSI-20)
  - data (LABSI-20)
  - processes (LABSI-20)
    - DataCuaca\_Testing (LABSI-20 - v1)
    - DataCuaca\_Training (LABSI-20 - v1)

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

ExampleSet (7 examples, 3 special attributes, 4 regular attributes)

Page 10 of 10

11:12 AM 10/10/2019

**RapidMiner Studio Trial 9.3.001 @ LABSI-20-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-20)
  - Connections (LABSI-20)
  - data (LABSI-20)
  - processes (LABSI-20)
  - DataCuaca\_Testing (LABSI-20 - v1, 10/10/19 11:08)
  - DataCuaca\_Training (LABSI-20 - v1, 10/10/19 11:08)

**ExampleSet (//Local Repository/DataCuaca\_Testing)**

Result History

Name	Type	Missing	Statistics	Filter (7 / 7 attributes):	Search for Attributes
Prediction <b>prediction(Berman_Tenis)</b>	Binominal	0	Least TIDAK (2)	Most YA (5)	Values YA (5), TIDAK (2)
Confidence_TIDAK <b>confidence(TIDAK)</b>	Real	0	Min 0.007	Max 0.856	Average 0.353
Confidence_YA <b>confidence(YA)</b>	Real	0	Min 0.144	Max 0.993	Average 0.647
<b>Cuaca</b>	Polynomial	0	Least Mendung (2)	Most Cerah (3)	Values Cerah (3), Mendung (2)
<b>Suhu</b>	Integer	0	Min 64	Max 83	Average 72.143
<b>Kelembaban_udara</b>	Integer	0	Min 65	Max 96	Average 79.571
<b>Berangin</b>	Polynomial	0	Least TIDAK (3)	Most YA (4)	Values YA (4), TIDAK (3)

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

11:12 AM 10/10/2019

## TUGAS

**Visual Studio Code**

File Edit Selection View Go Debug Terminal Help

Explorer

OPEN EDITORS

You have not yet opened a folder.

Open Folder

OUTLINE

The active editor cannot provide outline information.

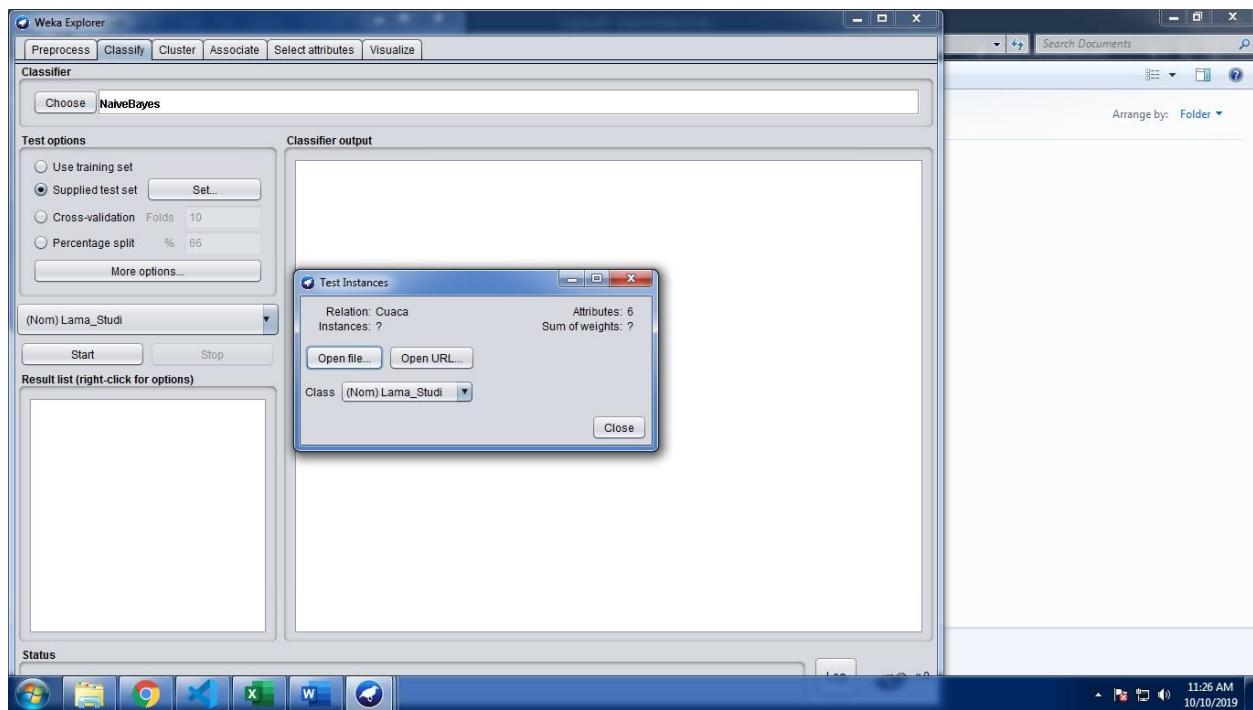
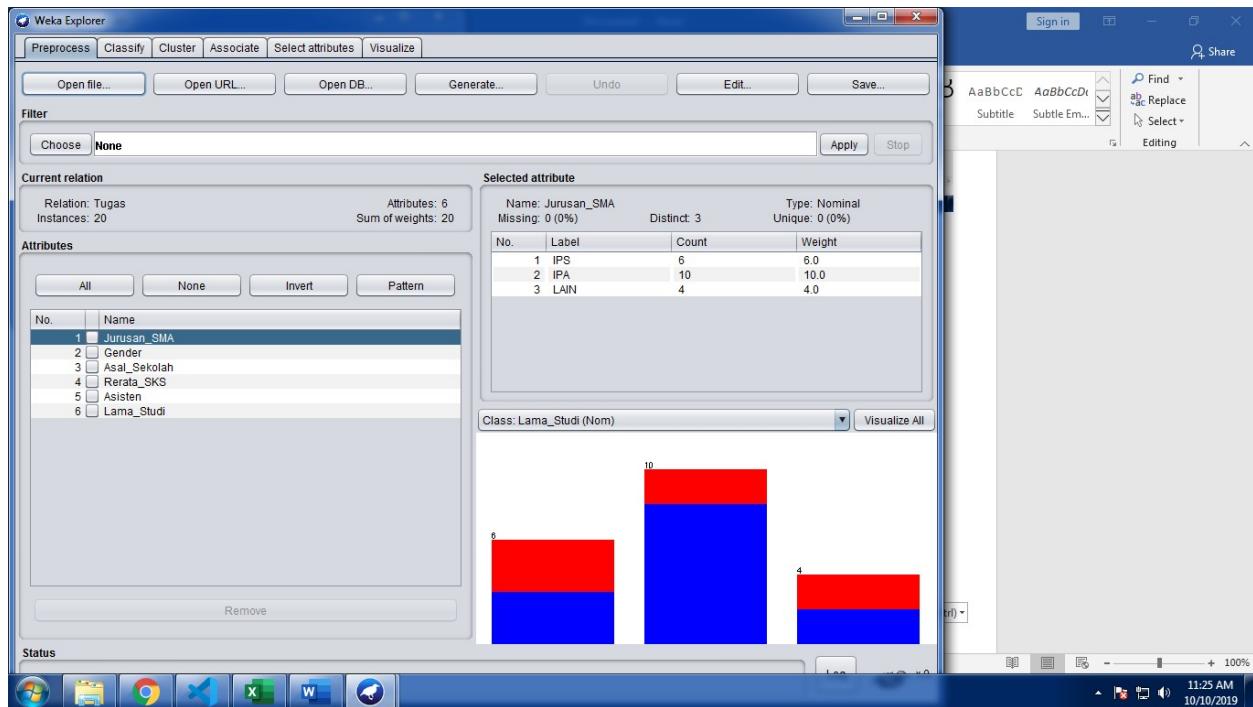
```

lcome cuaca.arff CucaTesting.arff (deleted) DATATestingSMAarff Tugas.arff DATATrainingSMAarff
1 @relation Tugas
2
3 @attribute Jurusan_SMA{IPS, IPA, LAIN}
4 @attribute Gender{WANITA, PRIA}
5 @attribute Asal_Sekolah{LUAR, SURAKARTA}
6 @attribute Rekata_SKS real
7 @attribute Asisten{TIDAK, YA}
8 @attribute Lama_Studi{TEPAT, TERLAMBAT}
9
10 @data
11 IPS,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

```

Ln 30, Col 38 Spaces: 4 UTF-8 CRLF Plain Text

11:24 AM 10/10/2019



**Weka Explorer**

- [Preprocess](#)
- [Classify](#)
- [Cluster](#)
- [Associate](#)
- [Select attributes](#)
- [Visualize](#)

**Classifier**

Choose **NaiveBayes**

**Test options**

- Use training set
- Supplied test set [Set...](#)
- Cross-validation Folds: 10
- Percentage split %: 66

[More options...](#)

(Nom) **Lama\_Studi**

[Start](#) [Stop](#)

**Result list (right-click for options)**

- [11:30:20 - misc.InputMappedClassifier](#)

**Classifier output**

```

Naive Bayes Classifier

Attribute      Class
              TEPAT TERLAMBAT
              (0.64) (0.36)
=====
Jurusan_SMA
IPS           4.0    4.0
IPA           9.0    3.0
LAIN          3.0    3.0
[total]        16.0   10.0

Gender
WANITA        5.0    2.0
PRIA          10.0   7.0
[total]        15.0   9.0

Asal_Sekolah
LUAR          5.0    3.0
SURAKARTA    10.0   6.0
[total]        15.0   9.0

Rerata_SKS
mean          19.5385 17.8571
std. dev.     1.5988  0.9897
weight sum    13       7
precision     1         1

Asisten

```

**Status**

1, Col 16 Spaces: 4 UTF-8 CRLF Plain Text

11:30 AM 10/10/2019

**ARFF-Viewer - C:\Users\LABSI-20\Documents\TugasTesting.arff**

[File](#) [Edit](#) [View](#)

**TugasTesting.arff**

**Relation: Tugas**

No	1: Jurusan_SMA	2: Gender	3: Asal_sekolah	4: Rerata_SKS	5: Asisten	6: Lama_Studi
	Nominal	Nominal	Nominal	Numeric	Nominal	Nominal
1	LAIN	WANITA	SURAKARTA	18.0	TIDAK	
2	IPA	PRIA	SURAKARTA	19.0	YA	
3	LAIN	PRIA	SURAKARTA	19.0	TIDAK	
4	IPS	PRIA	LUAR	17.0	TIDAK	
5	LAIN	WANITA	SURAKARTA	17.0	TIDAK	
6	IPA	WANITA	LUAR	18.0	YA	
7	IPA	PRIA	SURAKARTA	18.0	TIDAK	
8	IPA	PRIA	SURAKARTA	19.0	TIDAK	
9	IPS	PRIA	LUAR	18.0	TIDAK	
10	LAIN	WANITA	SURAKARTA	18.0	TIDAK	

**Find** **AaBbCcDd** **Subtitle** **Replace** **Select** **Editing**

rtf

11:32 AM 10/10/2019

## Data Testing

**Import Data - Select the cells to import.**

Select the cells to import.

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

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

← Previous → Next Cancel

Activate Wisdom of Crowds

**ExampleSet (/Local Repository/Tugas\_Testing)**

Views: Design Results Turbo Prep Auto Model

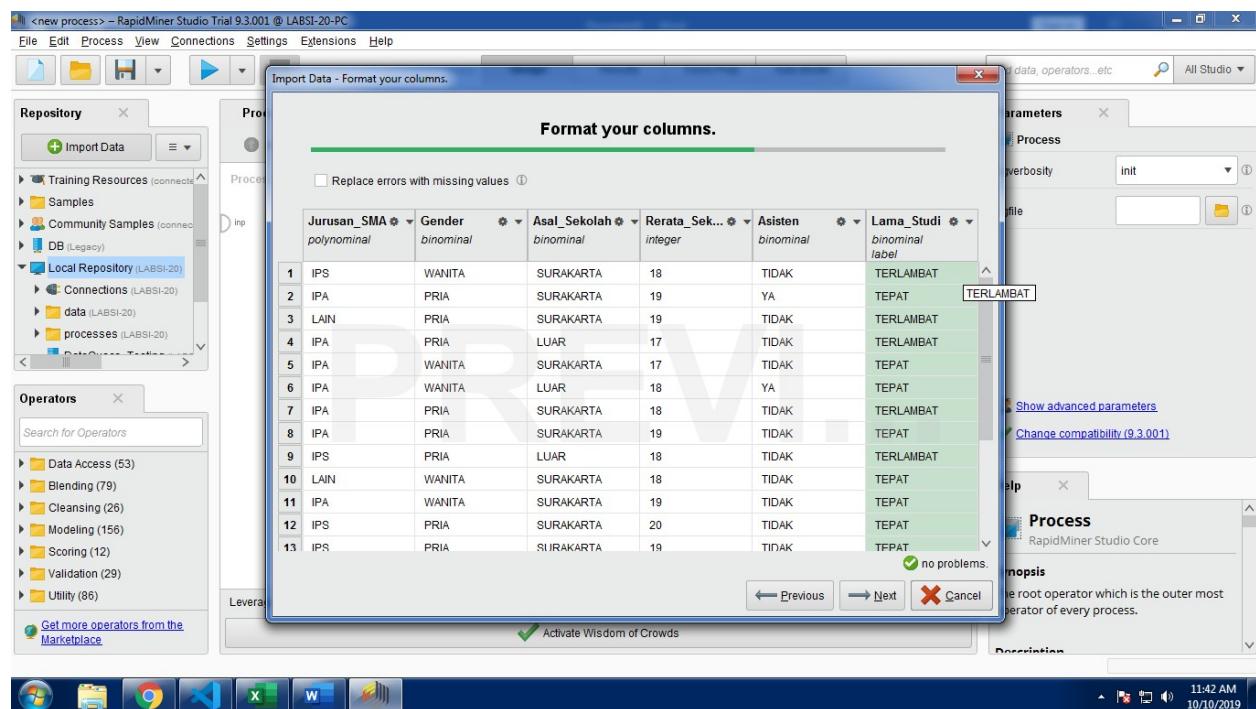
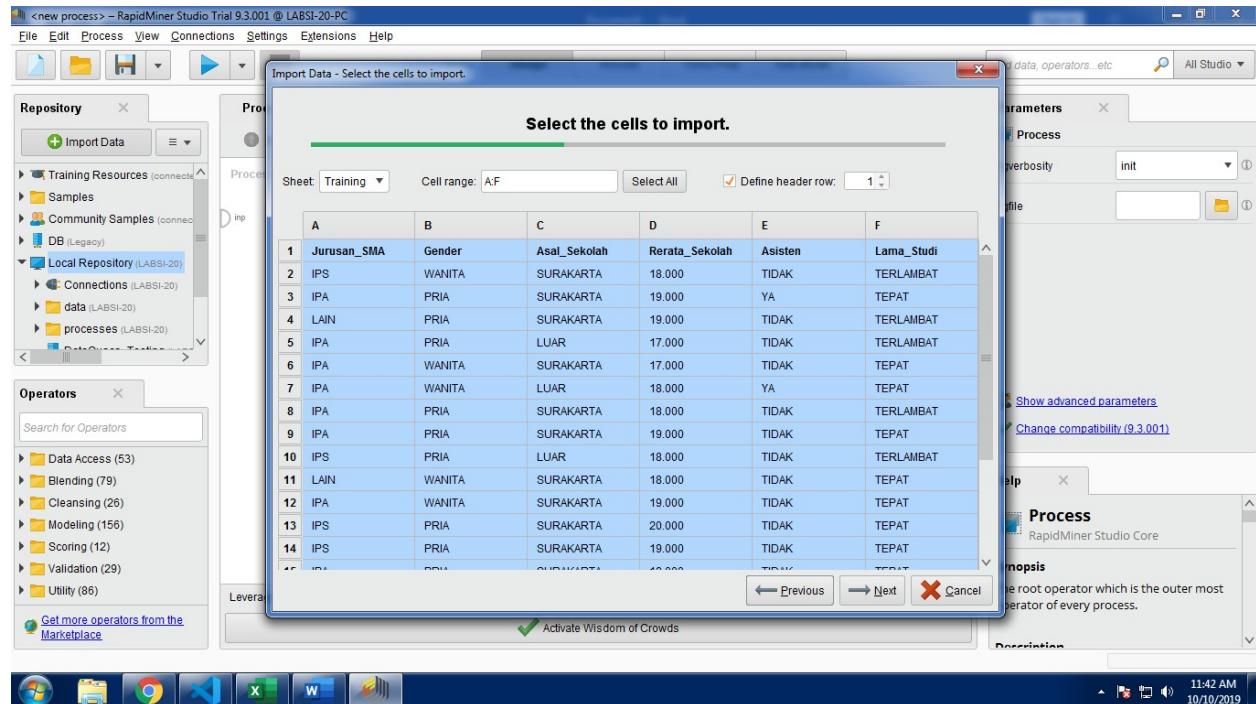
Repository

Row No.	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	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)

Calculating result: Visualizations

## Data Training



**<new process> – RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC**

File Edit Process View Connections Settings Extensions Help

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

Result History ExampleSet (/Local Repository/Tugas\_Training) ExampleSet (/Local Repository/Tugas\_Testing)

Data Statistics Visualizations Annotations

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

Row No.	Lama_Studi	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	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
15	TEPAT	IPA	PRIA	LUAR	22	YA

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

Import Data Training Resources (connected) Samples Community Samples (connected) DB (Legacy) Local Repository (LABSI-20) Connections (LABSI-20) data (LABSI-20) processes (LABSI-20) DataCuaca\_Testing (LABSI-20 - v1, 10/10/19 11:08) DataCuaca\_Training (LABSI-20 - v1, 10/10/19 11:08) Tugas\_Testing (LABSI-20 - v1, 10/10/19 11:41 AM) Tugas\_Training (LABSI-20 - v1, 10/10/19 11:43 AM)

**<new process\*> – RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC**

File Edit Process View Connections Settings Extensions Help

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

Repository Import Data Local Repository (LABSI-20) Connections (LABSI-20) data (LABSI-20) processes (LABSI-20) DataCuaca\_Testing (LABSI-20) DataCuaca\_Training (LABSI-20) Tugas\_Testing (LABSI-20) Tugas\_Training (LABSI-20)

Process

```

graph LR
    R1[Retrieve Tugas_Training] --> NB[Naive Bayes]
    R2[Retrieve Tugas_Testing] --> AM[Apply Model]
    NB --> AM
    AM --> R2
  
```

Operators Search for Operators Confidences (8) Apply Model Model Simulator Explain Predictions Prescriptive Analytics Validation (29) Get more operators from the marketplace

Parameters Process logverbosity: init logfile: Show advanced parameters Change compatibility (9.3.001)

Help Process RapidMiner Studio Core Synopsis The root operator which is the outer most operator of every process. Description

Activate Wisdom of Crowds

**RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC**

File Edit Process View Connections Settings Extensions Help

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

ExampleSet (/Local Repository/Tugas\_Training) ExampleSet (/Local Repository/Tugas\_Testing) ExampleSet (Apply Model)

Result History

Open in Turbo Prep Auto Model Filter (10 / 10 examples): all

Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_Sek...	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	0.868	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

ExampleSet (10 examples, 3 special attributes, 5 regular attributes)

Repository

- Training Resources (connected)
- Samples
- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-20)
  - Connections (LABSI-20)
  - data (LABSI-20)
  - processes (LABSI-20)
    - DataCuaca\_Testing (LABSI-20 - v1, 10/10/19 11:08)
    - DataCuaca\_Training (LABSI-20 - v1, 10/10/19 11:08)
    - Tugas\_Testing (LABSI-20 - v1, 10/10/19 11:41 AM)
    - Tugas\_Training (LABSI-20 - v1, 10/10/19 11:43 AM)

11:45 AM 10/10/2019

## Tambahan data testing

**RapidMiner Studio Trial 9.3.001 @ LABSI-20-PC**

File Edit Process View Connections Settings Extensions Help

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

Repository

Process

Parameters

Process

verbosity: init

logfile:

Operators

apply

Forecasting (1)

Scoring (2)

Confidences (1)

Apply Threshold

Apply Model

We found "Shapelet" and "Freemarker operator" in the Marketplace [Show me!](#)

Process

Naive Bayes

Apply Model

Help

Process

RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

Description

11:53 AM 10/10/2019

Kesimpulan :

Dewi dan Jono sama-sama lulus TEPAT.

## F. MODUL 9

### KEGIATAN PRAKTIKUM

#### Pohon Keputusan menggunakan Weka

[ ]

<b>Q Use training set</b>	<b>Kappa statistic</b>
Supplied test set	<b>Mean absolute error</b> 0
Set... _J	<b>Root mean squared error</b> 0
<b>Cross-validation</b>	<b>error</b> 0
<b>Folds ~ _J</b>	<b>Relative absolute error</b> 0
Percentage split % [66 ]	Root relative squared error 10 0
<b>Mort</b> <b>options..</b>	Coverage of ca.,, (0. 9S level) 80 14
(Nom) Be<maln_Tenls	Mean rel. rejoin size (0. 9S level)
--_sla__r_t,fr	Total Number!
Stop	Instances

### Detailed Accuracy By Class

	TP Rate	FP Rate	Precision	Recall	Specificity	MCC
Result first (right-click for options)	1.00	0.00	1.000	all	i , 000	i ,
	0	0			l , 000	
23-22j48					000	
	00	00	00	000	000	1.000
Weighted Avg.	00	0.0	00	000	i , 000	1.000
	00					

&:: Confusion Matrix==

a b <- classified. as

90 a=YA

0 5 b TIDAK

Status

OK

Web Explorer

Preprocess      Classify ~~~~~-----|||

Classifier

|    Choose ~48 -c 0.25 -M 2 -Classifier output

Test options

.. Use training  
set

S<4!)loed test set

Cuaca • Cerah

Set.

Kelembaban\_tklaro <• 7S: YA (2.0) [x]

Cross-validation

Kelembaban\_tklaro > 75: TIDAK (3.0)

Folds      10

Percentage split      %

66

CUaca • l'.endunq: YA (4 .0)

CUaca • HuJan

| (Norn)

I Beranqin • YA: TIDAK (2.0)

lennaon\_Tens

Beranqin • TIDAK: YA (3.0)

Sta  
rt

!IUllber

Resut &st (~t-  
dide for

o! Summary

options)

o! Leave

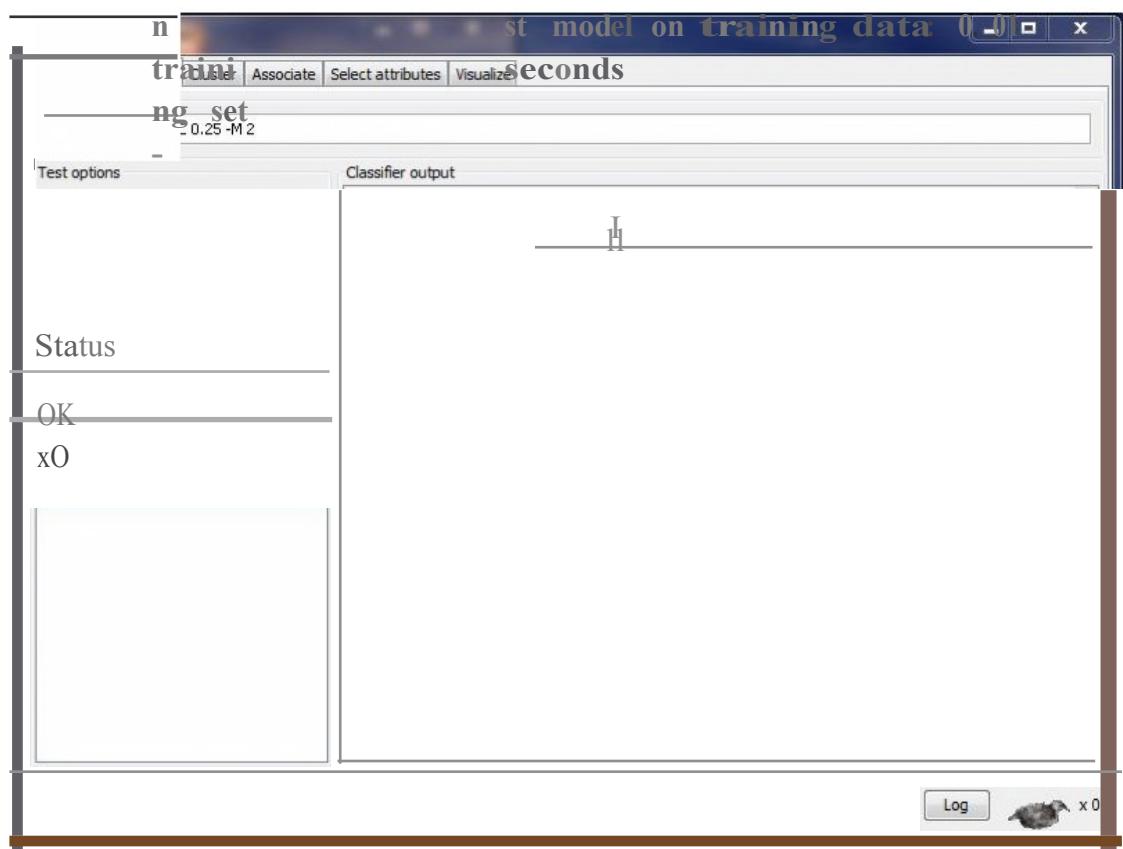
Size o!

23:22:2~

the tree

lrltel.>!8

old model: 0.01 aecnde



Weka  
Explorer

## Preprocess Classify

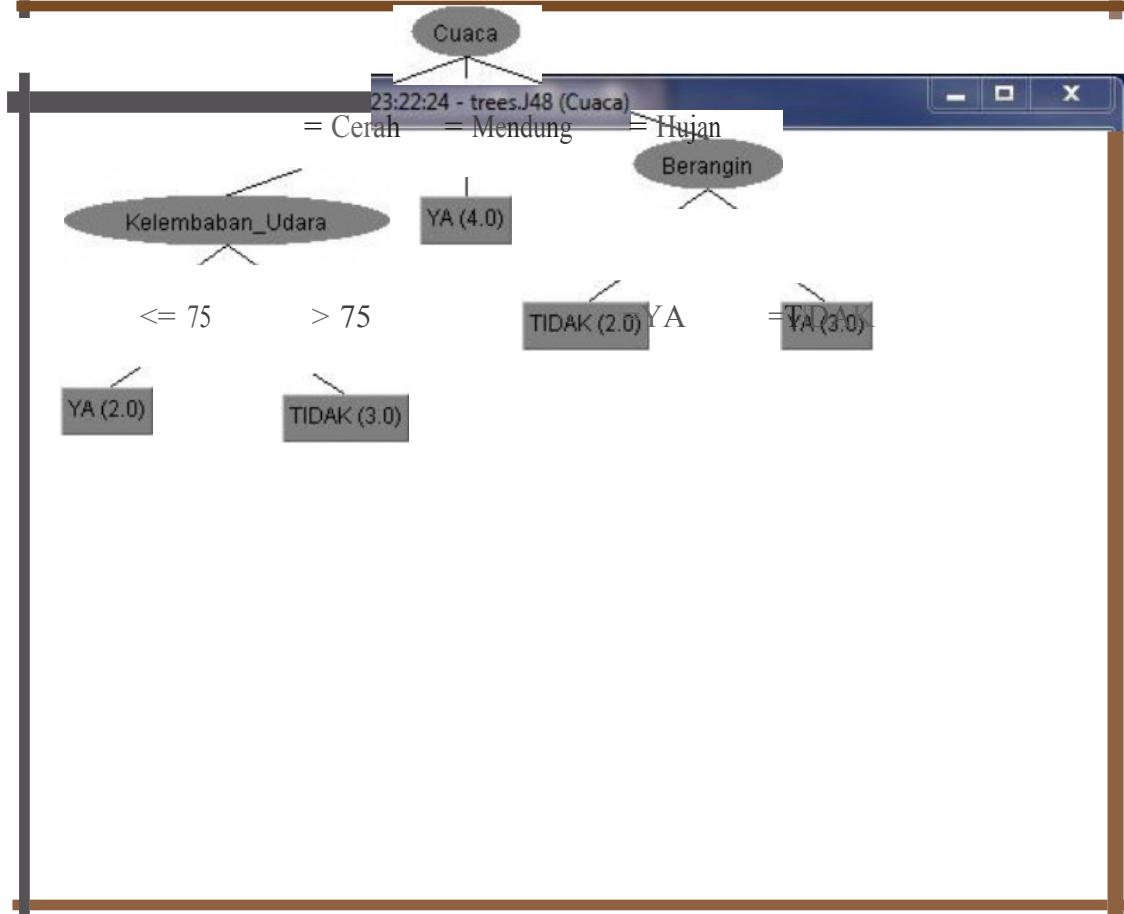
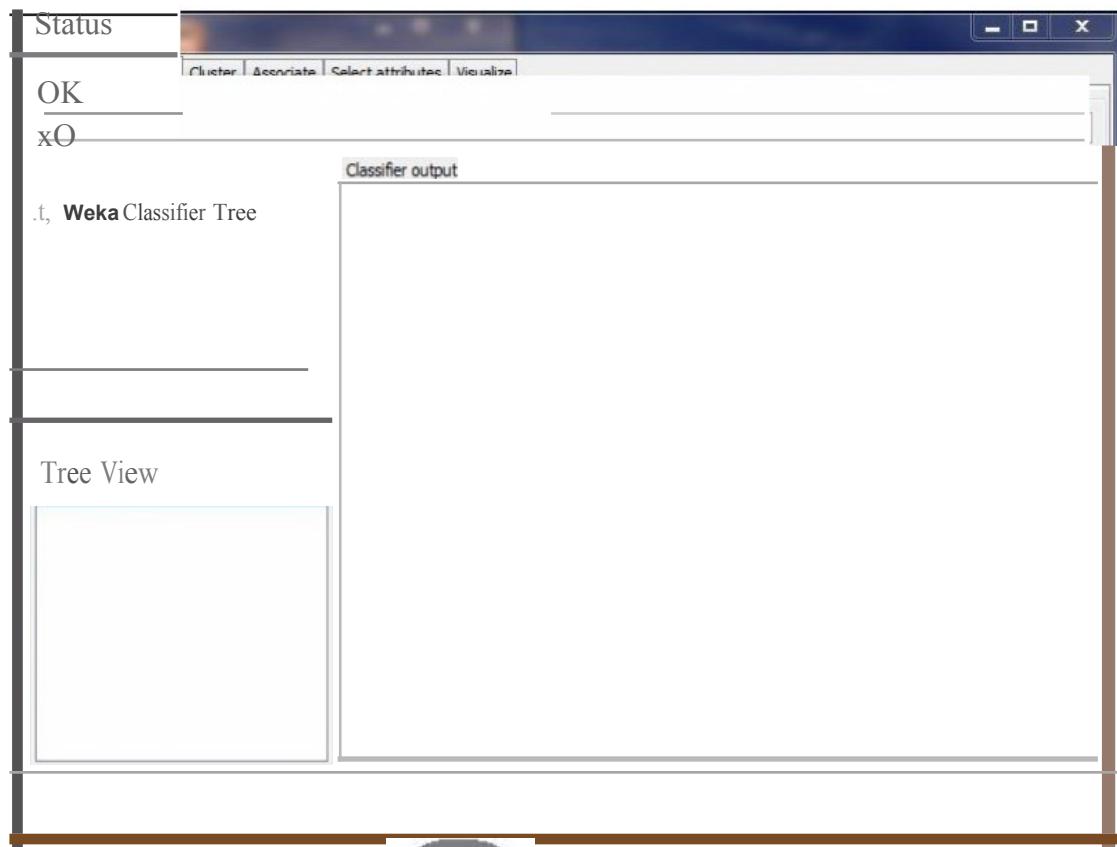
### Classifier

| Choose J48 -

### Test options

<input type="checkbox"/> use training set	- SWt:nar-y	
, Supplied test set	Classified Instance,	100
) Cross-validation Folds 10	Correctly Classified Instances,	14
Porcentage split 0%, 66	Incorrectly Classified Instances	0
	Kappastatistic	
	Mean absolute error	0
<u>options...</u> More	Root mean squared error	0
(Norn) Bermain_Tenis	Relative absolute error	0
	Root relative squared error	0
	Coverageof cases (0, 95 level)	100
<u>Start</u> - st -	Mean rel. region size (0, 95 level)	50
I   op	Total Number of Instances	H
Result list (right-click for options)	Detailed Accuracy By Class==	
23:22:24		
trees.J48		

	TP Rate	FP Rate	Precision	Recall	F-Measure	
Weighted Avg	1.000	0.000	1.000	1.000	1.000	1.000
	1.000	1.000	1.000	1.000	1.000	1.000
- Confusion Matrix -						
	b <-					
	classified ..					
	III					



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

- a. Simpul Akar = Cuaca
- b. Simpul internal = Kelembapan\_udara, dan Berangin
- c. Simpul Daun = YA, TIDAK

Klasifikasi yang terbentuk yaitu :

- a. Seseorang akan bermain tenis (YA) jika kondisi sebagai berikut :
  - i. Cuaca = Cerah, Kelembapan\_udara  $\leq 75$ . (nilai atribut lain diabaikan)
  - ii. Cuaca = Mendung. (kondisi lain diabaikan)
  - iii. Cuaca = Hujan, Berangin = TIDAK. (nilai atribut lain diabaikan)
- b. Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :
  - i. Cuaca = Cerah, Kelembapan\_udara  $> 75$ . (nilai atribut lain diabaikan)
  - ii. Cuaca = Hujan, Berangin = YA. (nilai atribut lain diabaikan)

## Pohon Keputusan menggunakan Rapid Miner

Result History    ExampleSet (/Local Repository/DataCuaca\_Training)    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)

Process    Cross Validation

Training    Testing

Operators

Parameters

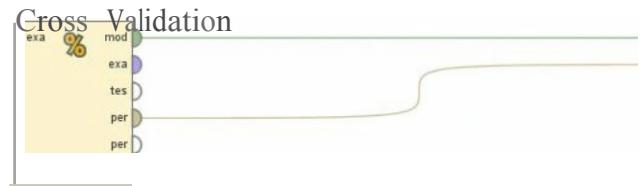
Help

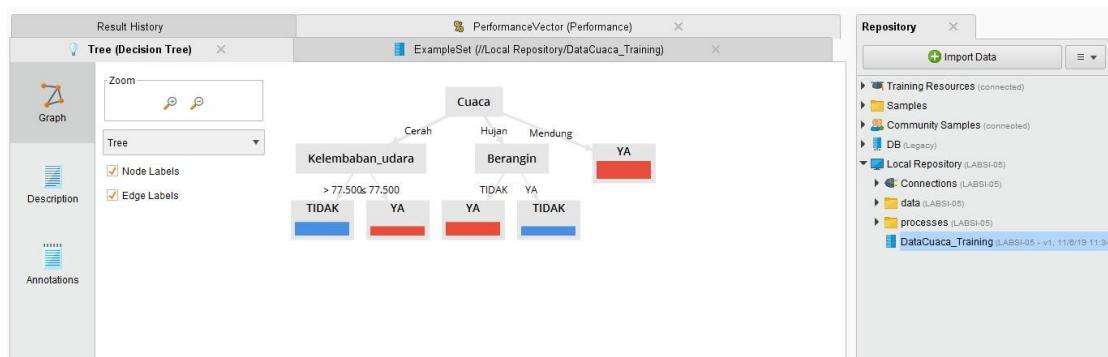
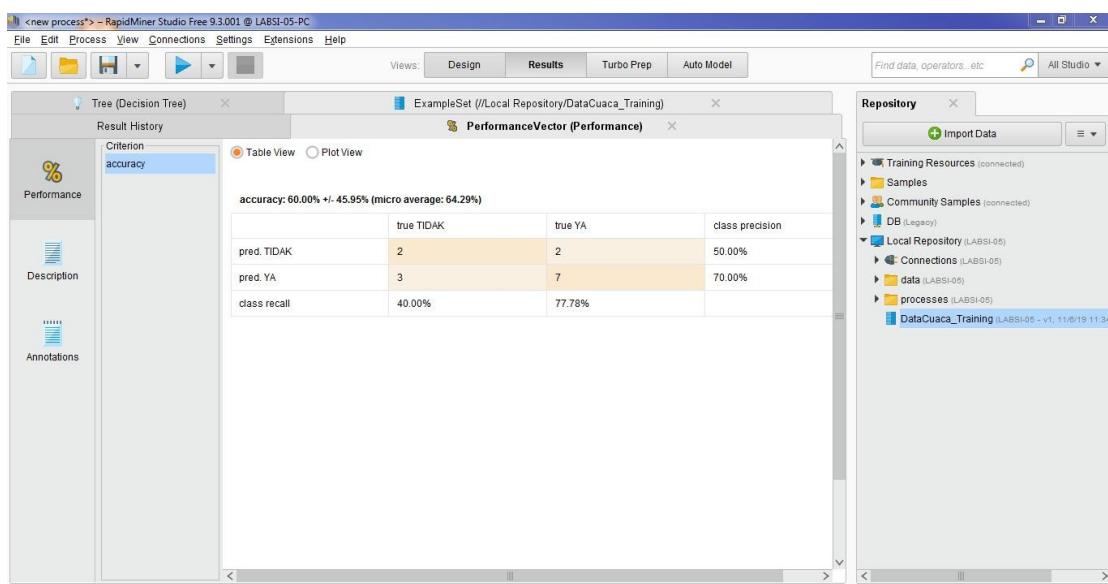
Process		Para mete rs
0 ~ ~	100% )i> Ji) Ji) + Ci... E3 •	attention information_g.. "
Cross Valdation	Decision Tree	
~ L_j	=GH2)	
	Apply Model PerfMmance	maximal depth 10
		./ apply pruning confidenc e 0.1
		./ apply prepruning
		minimal gain 0.01
		minimal leaf size
		!? Show advanced parameters
Process	Help	
• Process•	Decision Tree	
	• Concurrency	
	Tags;~~	
	~~fun	

100%fo      *ft j:)*      +  
~                  E3

P,

Retrieve  
DataCuaca\_  
...





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

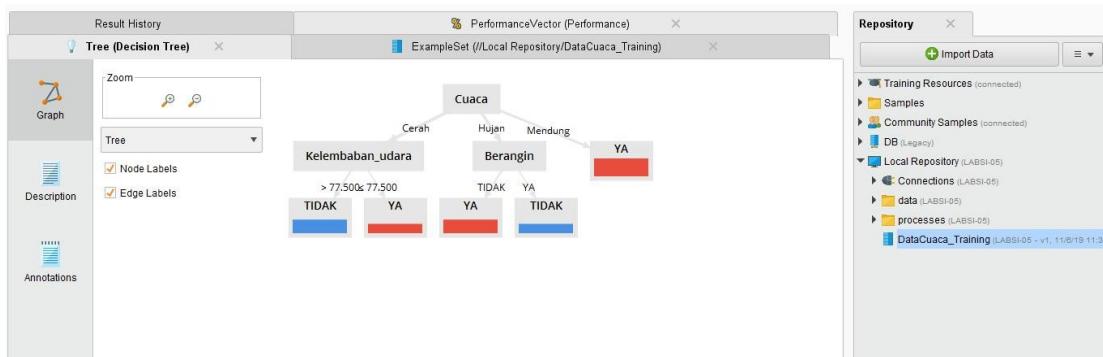
Klasifikasi yang terbentuk yaitu :

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

- i. Cuaca = Cerah, Kelembapan\_udara  $\leq 77,5$ . (nilai atribut lain diabaikan)
  - ii. Cuaca = Mendung. (kondisi lain diabaikan)
  - iii. Cuaca = Hujan, Berangin = TIDAK. (nilai atribut lain diabaikan)
- b. Seseorang tidak akan bermain tenis (TIDAK) jika kondisi sebagai berikut :
- i. Cuaca = Cerah, Kelembapan\_udara  $> 77,5$ . (nilai atribut lain diabaikan)
  - ii. Cuaca = Hujan, Berangin = YA. (nilai atribut lain diabaikan)

## TUGAS

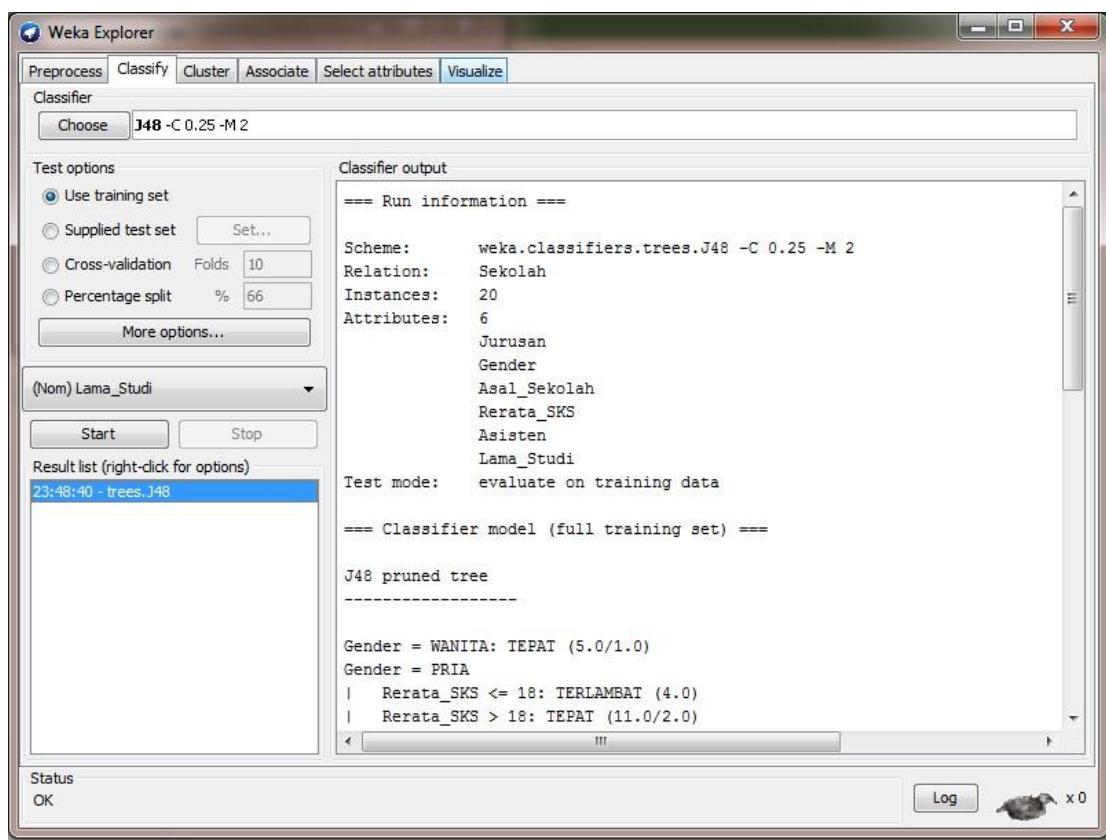
1. Berdasarkan pohon keputusan pada kegiatan 9.4.2 (menggunakan Rapid Miner), nilai kelas atribut Bermain\_Tenis pada tabel Testing sebagai berikut :



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

2. Berdasarkan pohon keputusan dari data Training pada Modul 7
  - i. Jumlah simpul daun pada pohon keputusan = 4
  - ii. Jumlah simpul keseluruhan pada pohon keputusan = 7
  - iii. Waktu yang dibutuhkan untuk proses pelatihan = 0 detik
  - iv. Tingkat ketepatan klasifikasi = 85%
  - v. Tingkat ketidaktepatan klasifikasi = 15%

### 3. Pohon keputusan berdasarkan data Sekolah.arff



## O WekaExp

lorer

### Preprocess

Classifier

Choose J48 -C 0.25 -M 2

G Use training set

J48 pruned tree

, Supplied test set SeL ..

, Cross-validation Folds 10

Percentage split % 66

under + WANITA: TEPAT {S.0/1.0}

under + PRIA

| Rerata\_SKS <= 18: TERLAMBAT (4.0)  
Rerata\_SKS > 18: TEPAT (11.0/2.0)

More options ..

Number of Leaves

(Nom) Lama\_Studi

Start

Stop>

51% of the tree

Result list (right-click for options)

23..qj:IO - trees..H6

Time taken to build model: 0 seconds

= Evaluation on training set =

Time taken to test model on training data: 0 seconds

====Summary==

Correctly Classified Instances

17

es

Incorrectly Classified Instances

3

es

III

Status

OK

x0

## O Weka Explorer

### Preprocess

Classifier

Choose J48 -C 0.25 -M 2

0 Use training set      Classify Cluster Associate Select attributes Visualize

Supplied test set Set ..

Test options	Classifier output	Correctly Classified Instances	17	85				
( Cross-validation Folds [ 10 ]	Incorrectly Classified Instances,							
<input type="radio"/> Percentage split % 66	Kappa statistic	3	15					
<a href="#">More options...</a>	Mean absolute error	0.6341						
	Root mean squared error	0.2436						
	Relative absolute error	0.349						
	Root relative squared error	53.0693	\					
	Coverage of cases (0.95 level)	73.1456	\					
	Mean rel. region size (0.95 level)	100						
	Total Number of Instances	90						
		20						
	Detailed Accuracy By Class							
		IP Rate	fP Rate	Precision	Recall	f Measure	HCC	
		1.000	0.429	0.813	1.000	0.897	0.681	
		0.571	0.000	1.000	0.571	0.727	0.681	
		Weighted Avg	0.850	0.279	0.878	0.850	0.837	0.681

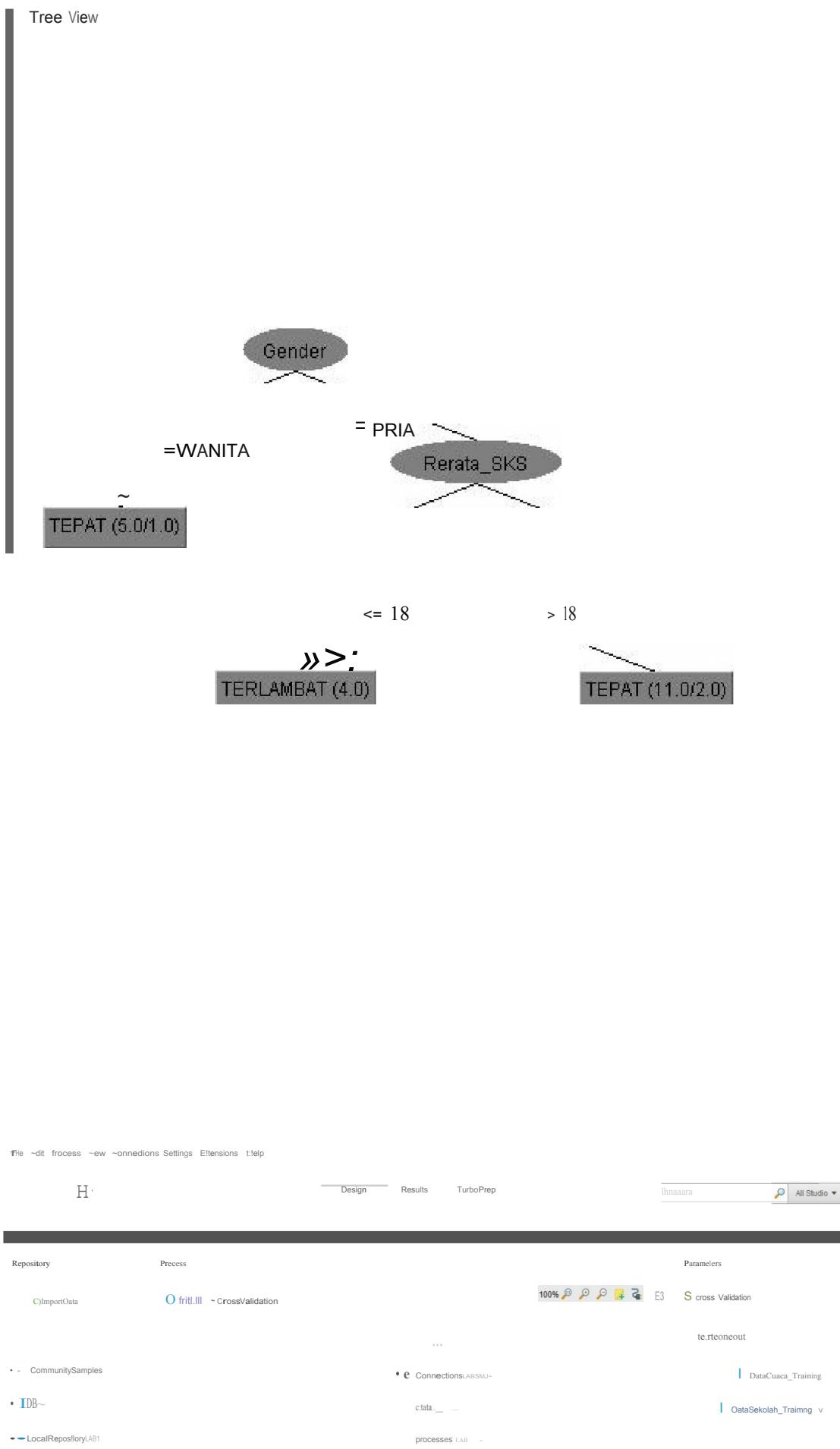
Result list (right-click for options)  
23.48:40 - trees J48

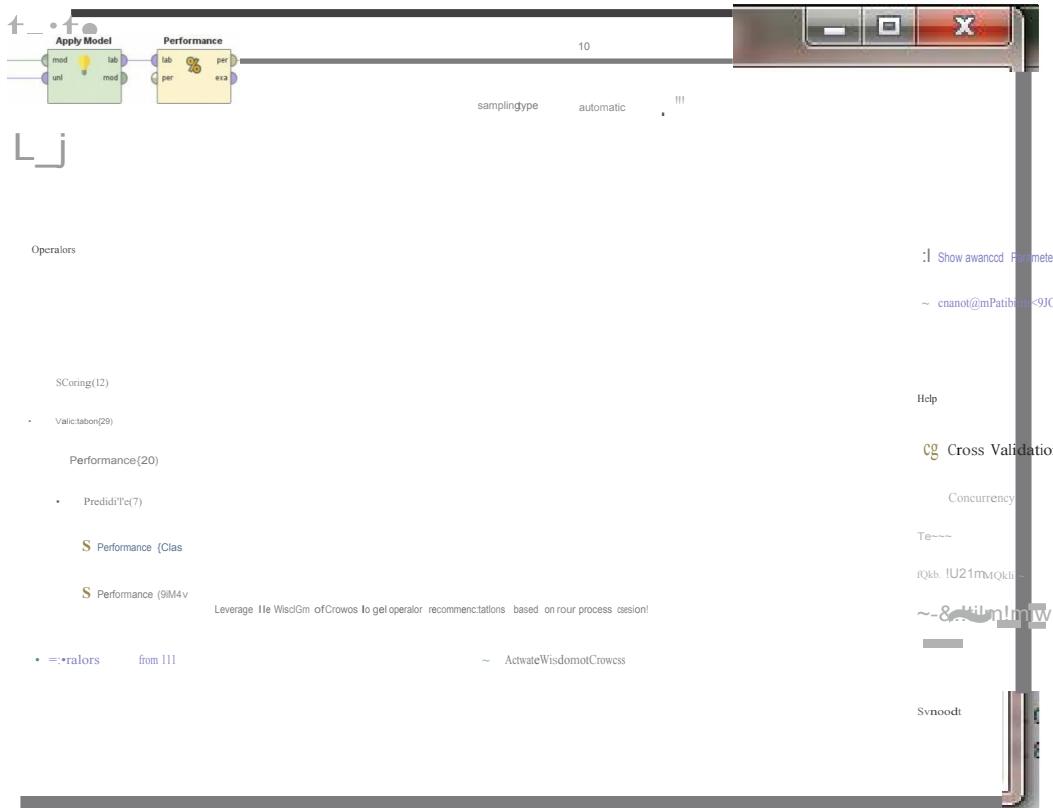
Cluster Associate Select attributes Visualize

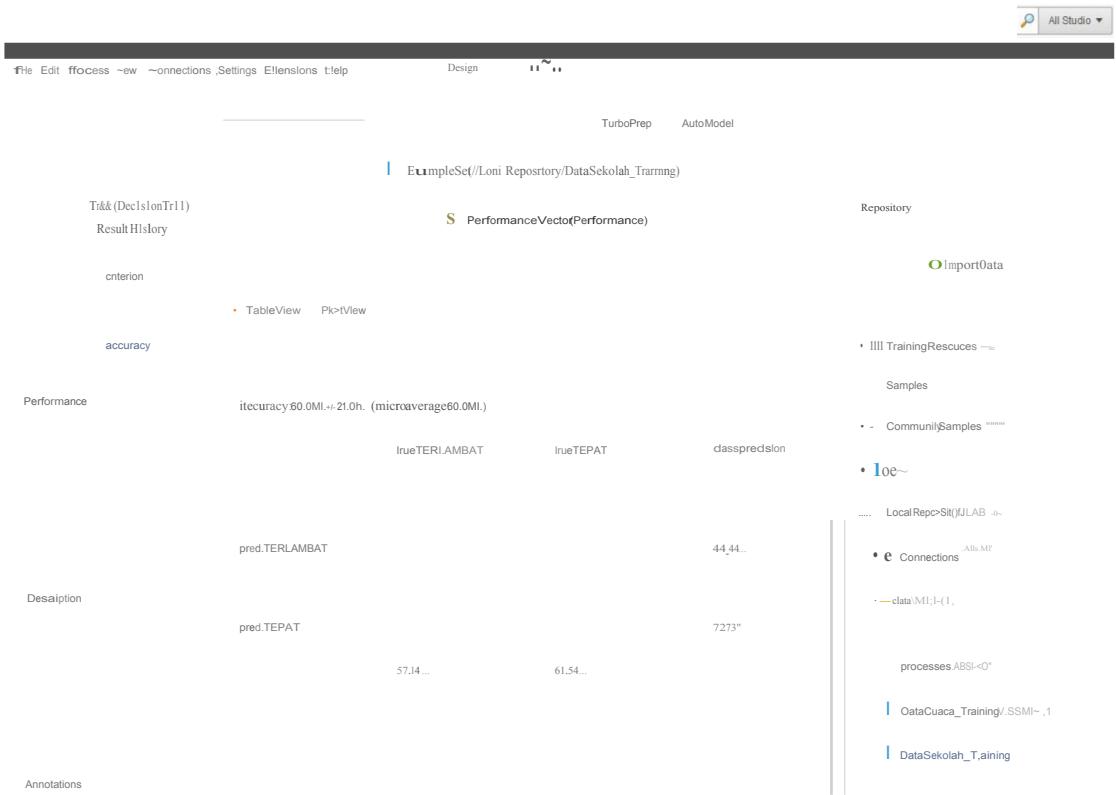
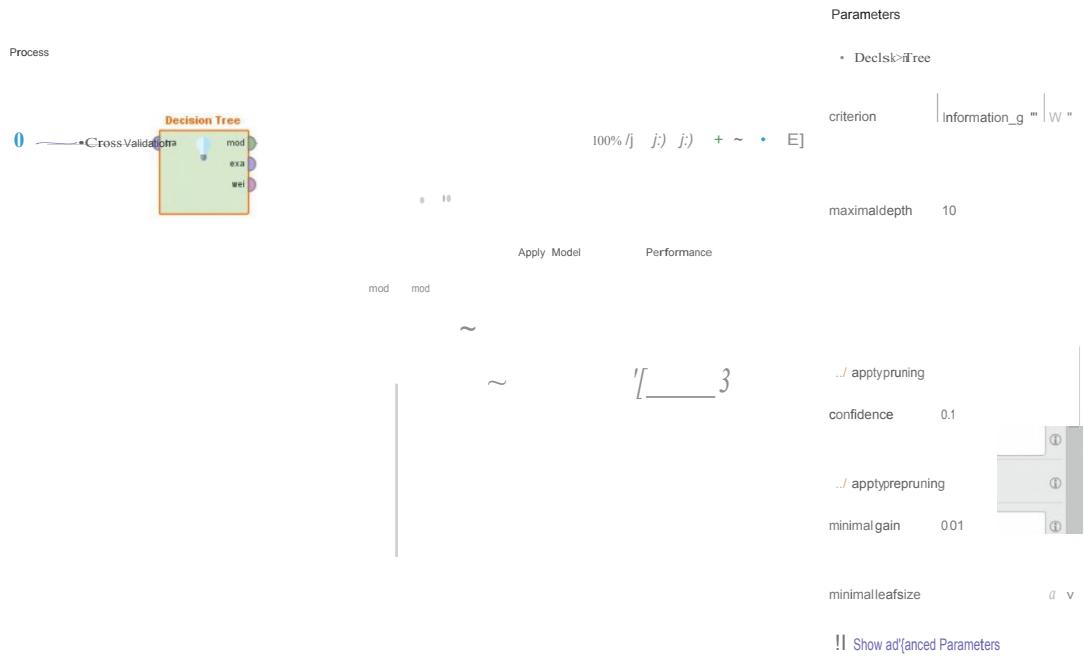
= Confusion Matrix=

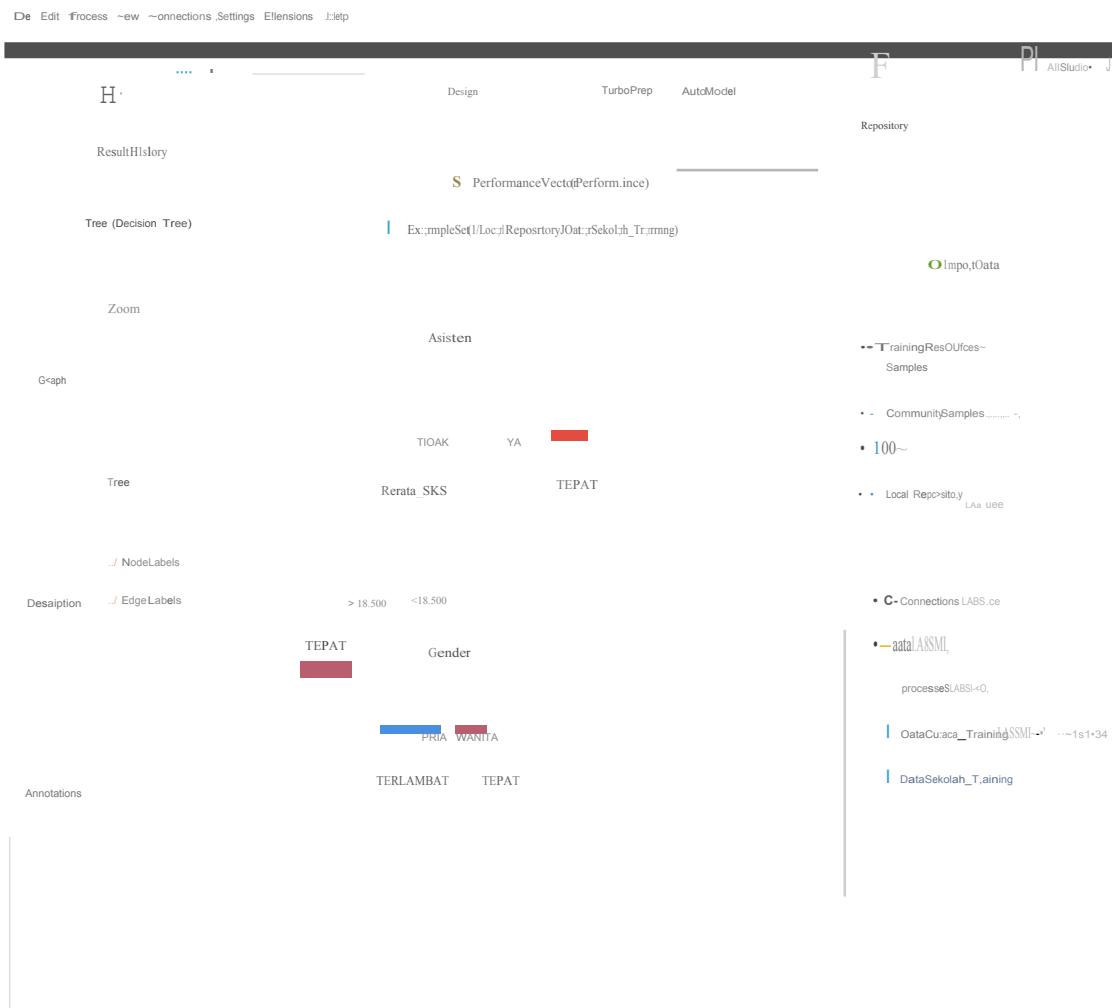
Test options	Classifier output
Status	b <-- classified as m
OK	x0

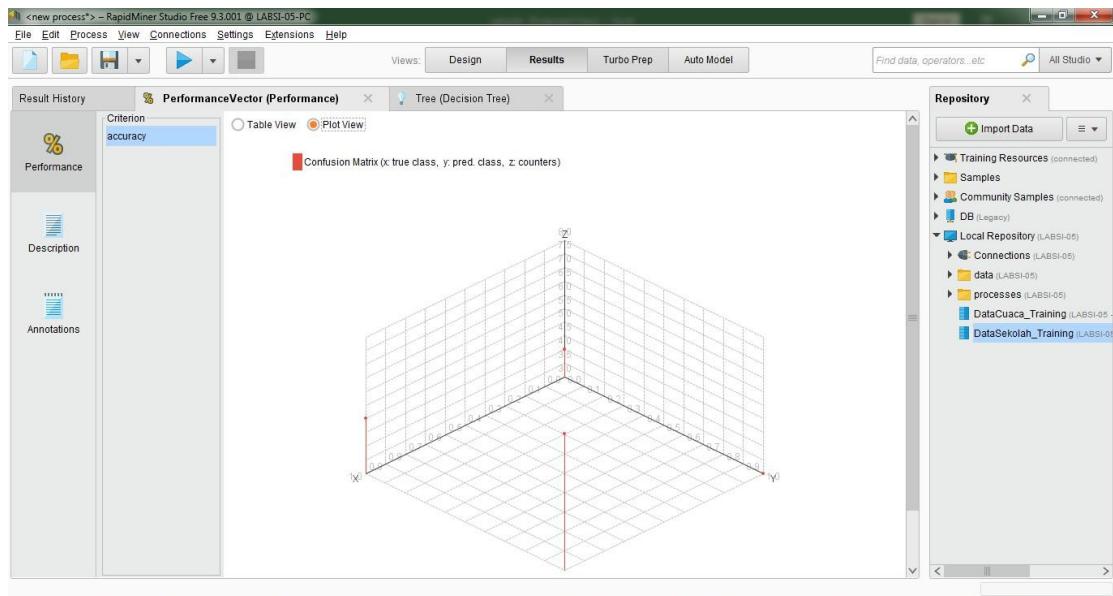
~ Weka Classifier Tree Visualizer. 23:48:40 - trees.J48 {Selcolah}









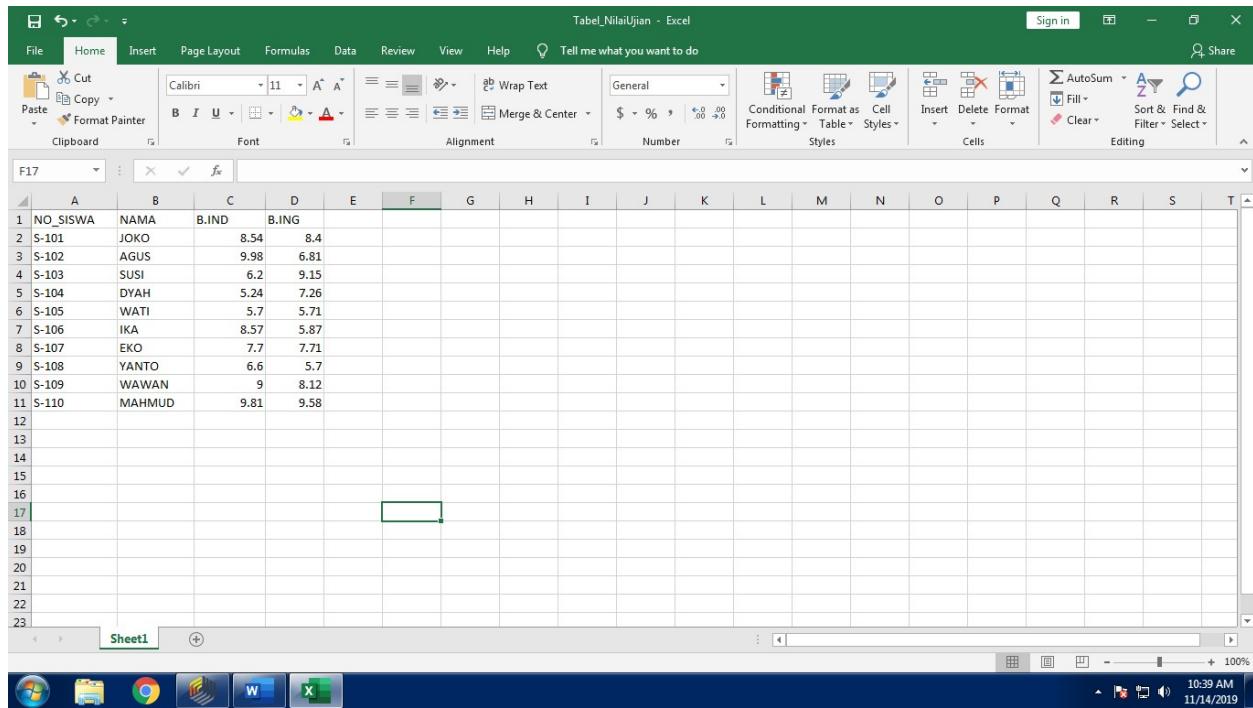


- i. Jumlah simpul daun pada pohon keputusan = 3
- ii. Jumlah simpul keseluruhan pada pohon keputusan = 5
- iii. Waktu yang dibutuhkan untuk proses pelatihan = 0 detik iv. Tingkat ketepatan klasifikasi = 85%
- v. Tingkat ketidaktepatan klasifikasi = 15%

## G. MODUL 10

### KEGIATAN PRAKTIKUM

- Berikut table data nilai siswa :

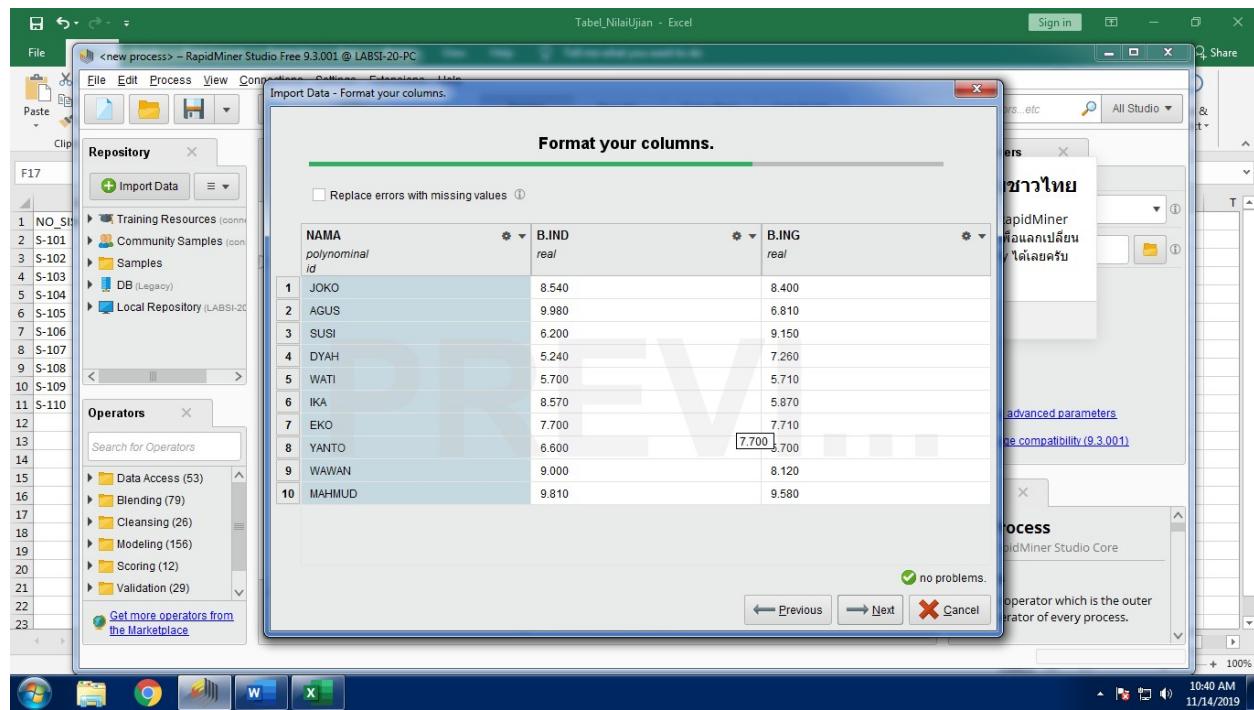
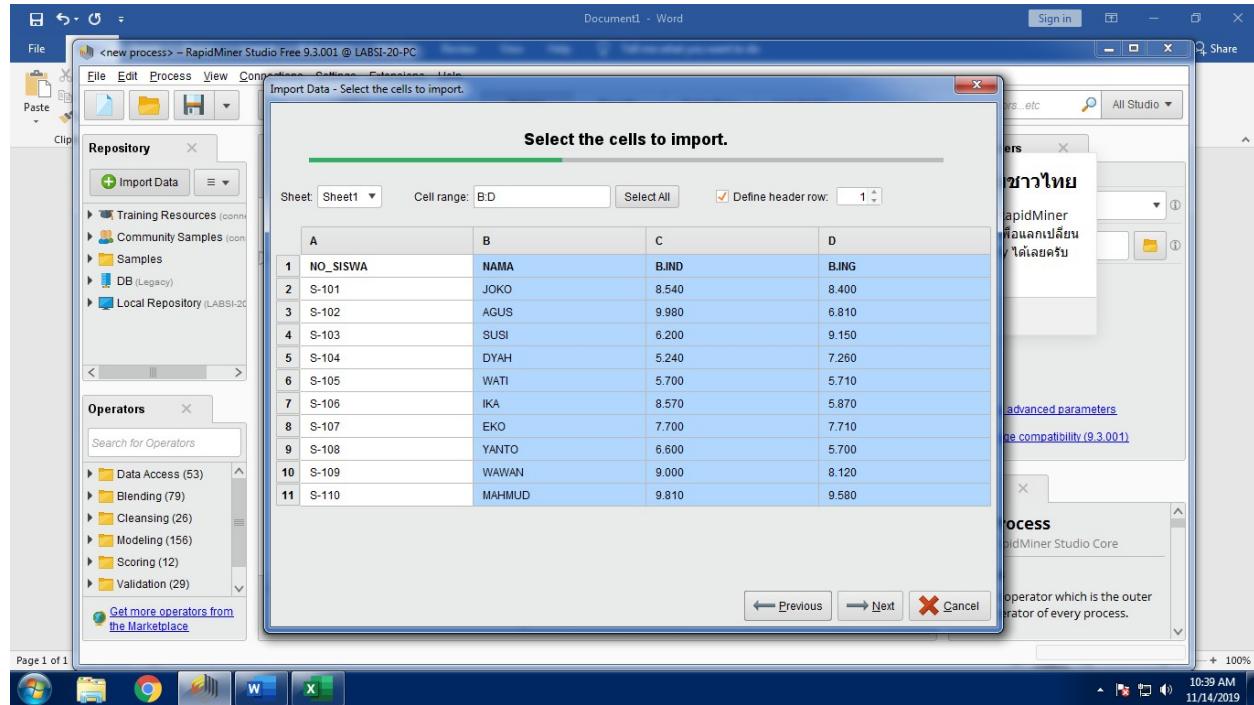


The screenshot shows a Microsoft Excel spreadsheet titled "Tabel\_NilaiUjian - Excel". The table has columns labeled NO\_SISWA, NAMA, B.IND, and B.ING. The data consists of 11 rows of student information. Row 17 is highlighted in green, indicating it is selected.

NO_SISWA	NAMA	B.IND	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

- Gunakan file Tabel\_NilaiUjian.xlsx sebagai data yang akan digunakan dalam proses Clustering.

Lalu import ke dalam aplikasi RapidMiner.



Result History

ExampleSet (/Local Repository/Data\_NilaiUjian)

Data

Row No.	NAMA	BJND	BJNG
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	8.120
10	MAHMUD	9.810	9.580

ExampleSet (10 examples, 1 special attribute, 2 regular attributes)

Page 2 of 2

10:41 AM 11/14/2019

- Tambahkan operator k-Means. Lalu Jalankan dengan menekan tombol run (F11)

Repository

Process

Process

Retrieve Data\_NilaiUjian → Clustering → SVD

Parameters

SVD (Singular Value Decomposition)

dimensionality reduction: fixed number

dimensions: 1

Operators

svd

Process

Help

Singular Value Decomposition

RapidMiner Studio Core

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

10:43 AM 11/14/2019

- Nilai Eigenvalue

The screenshot shows the RapidMiner Studio interface with the title bar <new process\*> – RapidMiner Studio Free 9.3.001 @ LABSI-20-PC. The main window displays the Results view for an SVD process. A table titled 'Eigenvalues' shows the following data:

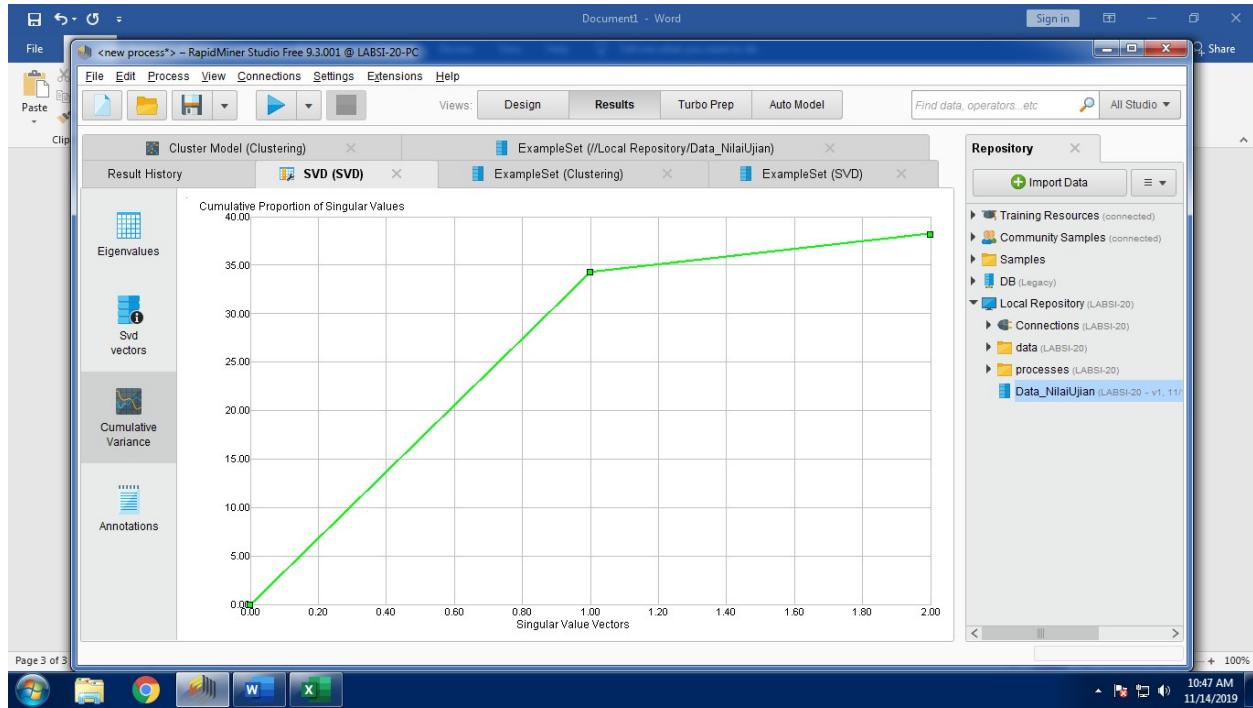
Component	Singular Value	Proportion of Singular V...	Cumulative Singular Val...	Cumulative Proportion o...
SVD 1	34.340	0.898	34.340	0.898
SVD 2	3.906	0.102	38.246	1.000

- Nilai Svd vectors

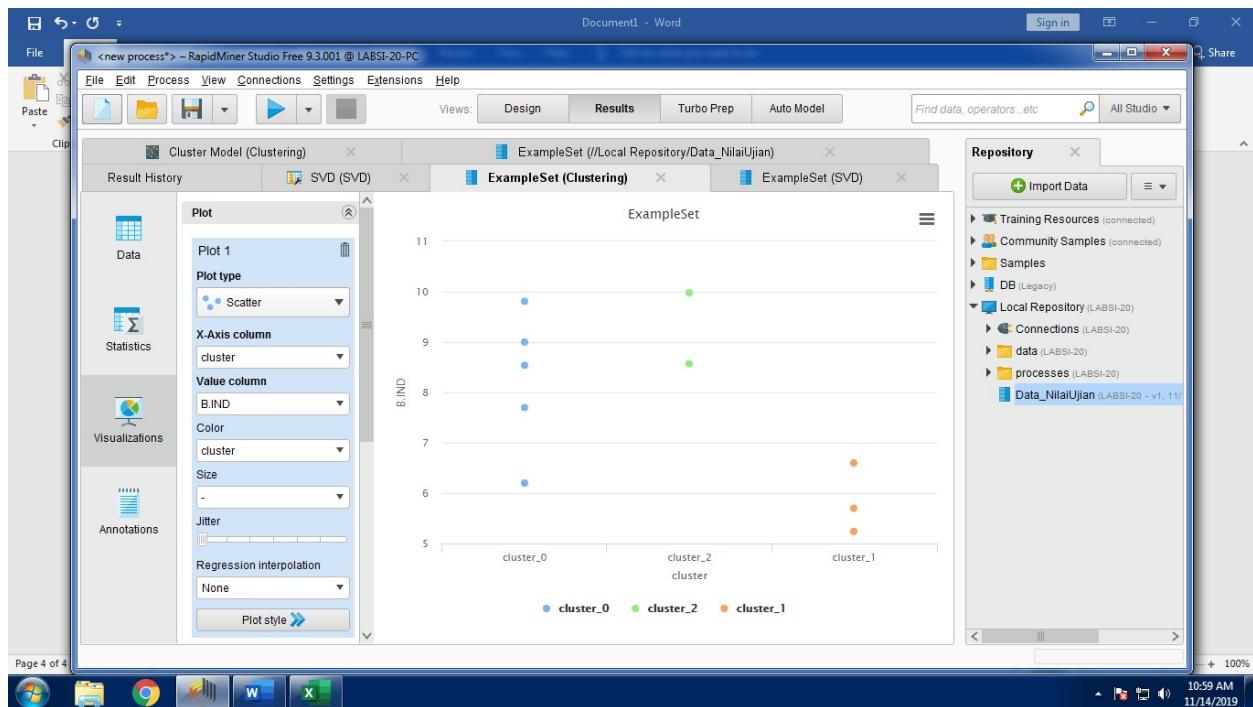
The screenshot shows the RapidMiner Studio interface with the title bar <new process\*> – RapidMiner Studio Free 9.3.001 @ LABSI-20-PC. The main window displays the Results view for an SVD process. A table titled 'Svd vectors' shows the following data:

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

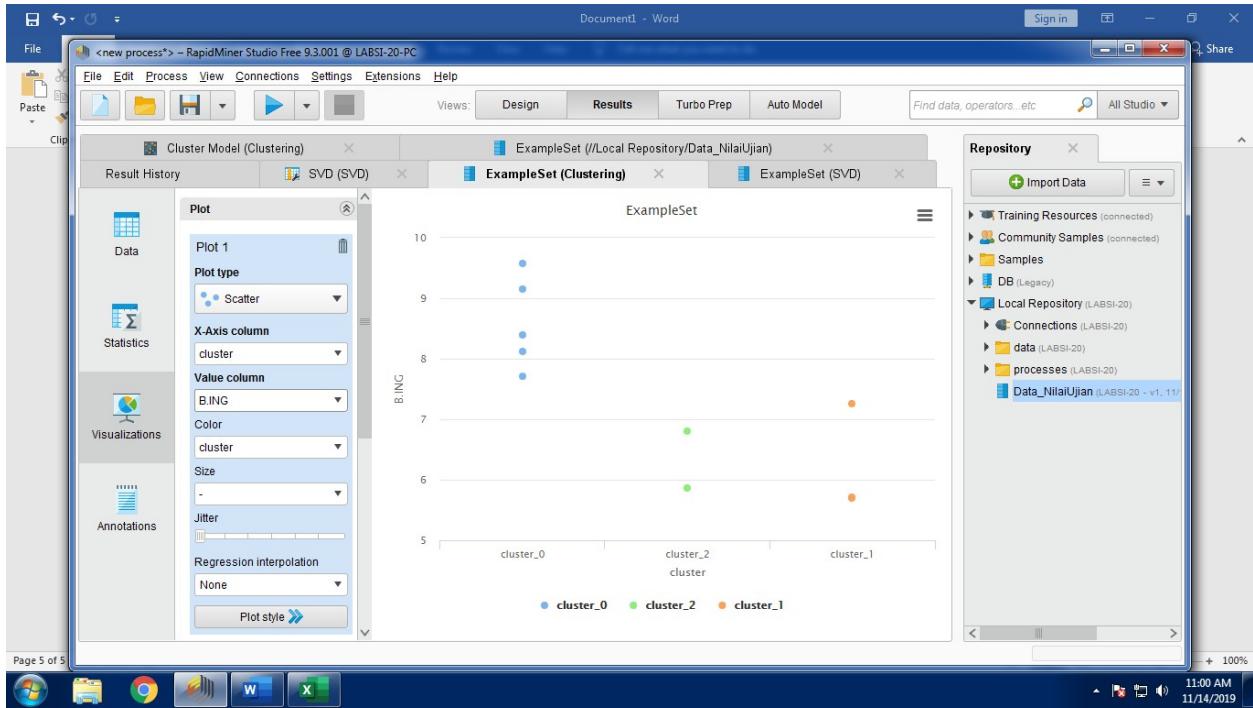
- Nilai Cumulative Variance



- Kelompok siswa bidang B.indonesia



- Kelompok siswa bidang B.inggris



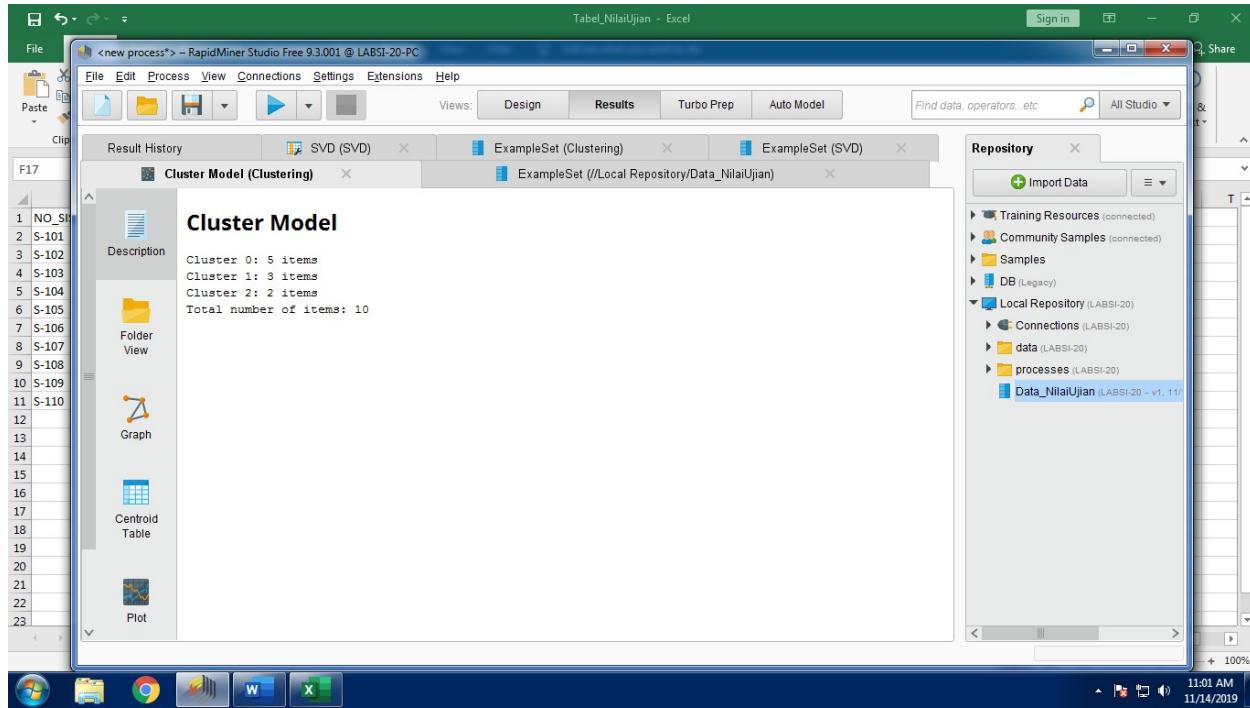
- Kelompok masing-masing siswa yang dikelompokkan bedasarkan cluster 0, cluster 1, cluster 2.

The screenshot shows the RapidMiner Studio interface. In the center, there is a table titled "ExampleSet (10 examples, 2 special attributes, 2 regular attributes)". The columns are "Row No.", "NAMA", "cluster", "BJND", and "B.ING". The data is as follows:

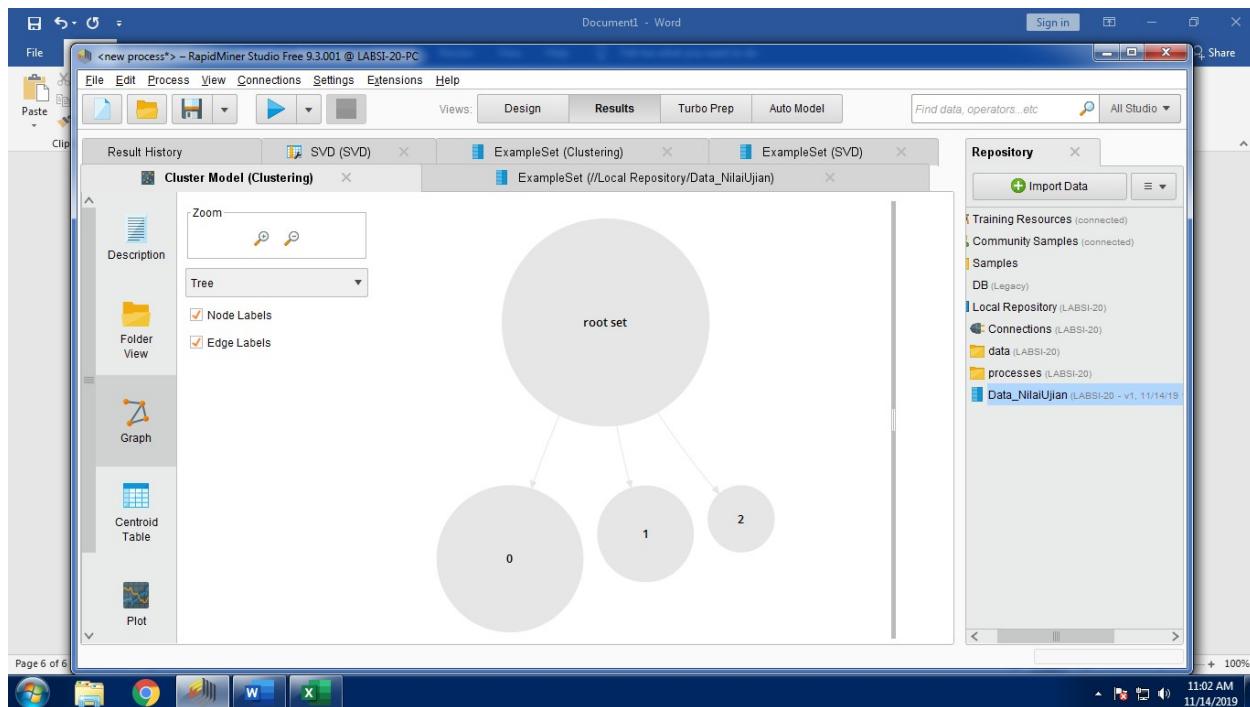
Row No.	NAMA	cluster	BJND	B.ING
1	JOKO	cluster_0	8.540	8.400
2	AGUS	cluster_2	9.980	6.810
3	SUSI	cluster_0	6.200	9.150
4	DYAH	cluster_1	5.240	7.260
5	WATI	cluster_1	5.700	5.710
6	IKA	cluster_2	8.570	5.870
7	EKO	cluster_0	7.700	7.710
8	YANTO	cluster_1	6.600	5.700
9	WAWAN	cluster_0	9	8.120
10	MAHMUD	cluster_0	9.810	9.580

The left sidebar shows the "Data" tab is selected. The right sidebar shows the "Repository" tab with a tree view of training resources, community samples, samples, DB (Legacy), and Local Repository (LABSI-20) containing connections, data, processes, and the current dataset "Data\_NilaiUjian".

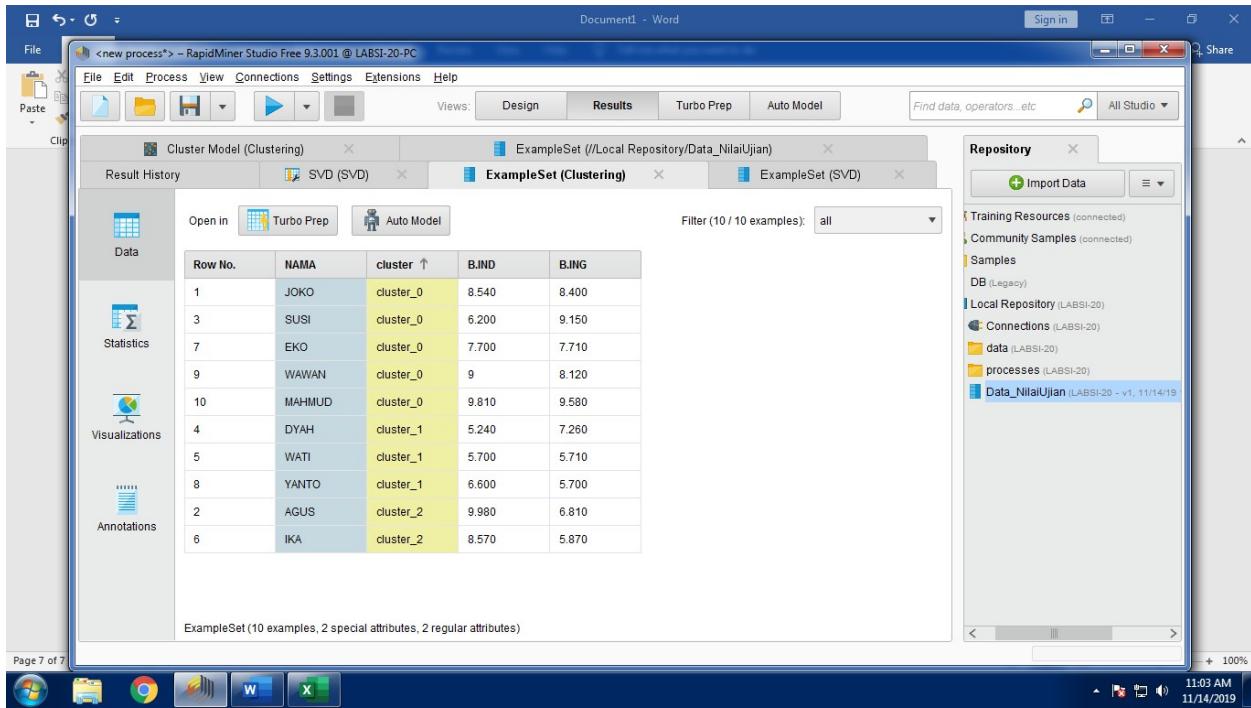
- Description



- Graph



## Kesimpulan :



1. Cluster 2 yang diajukan untuk lomba olimpiade bidang B.Indonesia
2. Cluster 0 yang diajukan untuk lomba olimpiade bidang B.Inggris

## Tugas

- Berikut adalah table siswa dan memasukkan nilai secara random dengan =5+RAND()\*5

NO_SISWA	NAMA	B.IND	B.ING	MTK	IPA
S-101	JOKO	6.33	7.03	9.95	9.00
S-102	AGUS	5.67	5.28	6.67	8.24
S-103	SUSI	5.69	7.92	5.90	5.17
S-104	DYAH	6.67	5.57	9.11	8.71
S-105	WATI	6.04	7.82	5.11	9.95
S-106	IKA	7.66	6.07	6.31	6.97
S-107	EKO	5.48	8.77	5.28	9.66
S-108	VANDO	3.44	9.53	7.62	3.80
S-109	WANWAN	5.54	5.31	5.26	5.15
S-110	MAMNUD	9.18	5.88	6.42	5.13
S-111	BUDI	8.15	5.18	6.86	5.15
S-112	SANTI	8.98	7.28	8.51	9.54
S-113	DIAN	6.10	7.66	9.33	9.23
S-114	DANI	5.17	7.34	6.86	3.40
S-115	AHMAD	5.14	8.93	6.63	6.49
S-116	BAYU	5.06	6.27	7.29	8.96
S-117	RISA	3.85	9.61	6.86	7.61
S-118	RANI	5.13	6.31	3.83	6.07
S-119	YAHYA	7.55	9.88	6.92	5.51
S-120	PARTH	5.78	5.24	9.88	8.83
S-121	INDAH	8.90	5.17	3.08	3.00
S-122	JOONO	5.83	8.24	5.58	5.83
S-123	SARAH	6.10	7.05	6.81	7.45
S-124	RAMA	5.13	5.26	6.41	5.10
S-125	BAMBANG	5.52	6.33	5.30	9.63
S-126	HADI	6.27	8.67	6.98	9.15
S-127	NANA	7.98	7.69	7.68	9.34
S-128	FEBRI	6.75	5.80	7.87	5.76
S-129	DENI	5.85	5.97	7.12	6.04
S-130	TONI	6.01	5.64	5.38	6.10

- Gunakan file Tugas\_NilaiUjian.xlsx sebagai data yang akan digunakan dalam proses Clustering.  
Lalu import ke dalam aplikasi RapidMiner.

kegiatanmodul10\_dwdm - Word

<new process\*> - RapidMiner Studio Free 9.3.001 @ LABSI-20-PC

File Edit Process View Connections Options Extensions Help

Import Data - Select the cells to import.

Select the cells to import.

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

A	B	C	D	E	F
13	S-112	SANTI	8.975	7.280	8.514
14	S-113	DIAN	6.098	7.660	9.331
15	S-114	DANI	5.170	7.338	6.863
16	S-115	AHMAD	5.140	8.933	6.625
17	S-116	BAYU	5.062	6.268	7.287
18	S-117	RISA	9.846	9.611	6.862
19	S-118	RANI	5.126	6.906	9.694
20	S-119	YANI	7.547	9.881	6.825
21	S-120	RATIH	9.785	5.239	9.893
22	S-121	INDAH	8.901	5.171	9.076
23	S-122	JONO	9.832	8.238	5.557
24	S-123	SARAH	6.105	7.050	6.814
25	S-124	RAMA	5.127	5.262	6.411
26	S-125	BAMBANG	5.523	6.326	5.300

← Previous → Next ✖ Cancel

Repository Operators

Import Data

Samples DB (Legacy) Local Repository (LABSI-20) Connections (LABSI-20) data (LABSI-20) processes (LABSI-20) Data\_NilaiUjian (LABSI-20)

svd

Cleansing (1)

Dimensionality Reduction Singular Value Decomposition

No results were found.

Page 8 of 8

11:13 AM 11/14/2019

100%

The screenshot shows the RapidMiner Studio interface with a data import dialog open. The dialog title is "Import Data - Select the cells to import." It contains a table with 14 rows and 6 columns (A-F). The first column (A) contains student IDs like S-112 to S-125. The second column (B) contains names like SANTI, DIAN, DANI, AHMAD, etc. Columns C-F contain numerical values. A checkbox "Define header row" is checked, and the value "1" is selected. Below the table are "Previous" and "Next" buttons, and a "Cancel" button. The background shows the RapidMiner studio environment with various operators and repositories visible.

Screenshot of RapidMiner Studio Free 9.3.001 showing the 'Format your columns' dialog box. The dialog shows a table with 13 rows and 6 columns: NAMA, B.IND, B.ING, MTK, IPA. The last column, IPA, has a value of 7.619 highlighted. The dialog includes a 'no problems.' message and buttons for Previous, Next, and Cancel.

NAMA	B.IND	B.ING	MTK	IPA
1 JOKO	6.330	7.029	9.950	9.001
2 AGUS	5.869	5.275	6.666	8.243
3 SUSI	5.692	7.916	5.903	5.166
4 DYAH	6.671	5.570	9.109	8.714
5 WATI	6.039	7.820	5.111	9.961
6 IKA	7.658	6.069	6.309	6.972
7 EKO	5.485	8.767	5.277	9.657
8 YANTO	9.438	9.526	7.619	9.803
9 WAWAN	5.839	5.907	7.619	5.454
10 MAHMUD	9.182	5.976	6.424	8.129
11 BUDI	8.155	5.177	6.863	5.152
12 SANTI	8.975	7.280	8.514	9.545
13 DIAN	6.098	7.660	9.331	9.226

- Tambahkan operator k-Means. Lalu Jalankan dengan menekan tombol run (F11)

Screenshot of RapidMiner Studio Free 9.3.001 showing the process editor. The process flow starts with 'Retrieve Data\_Tugas...' (Input) connected to a 'Clustering' operator. The output of 'Clustering' is connected to an 'SVD' operator. The 'Parameters' panel on the right is set for 'Clustering (k-Means)' with max runs: 10, determine good start values checked, measure types: Numerical, numerical measure: EuclideanDistance, max optimization: 100. The 'Help' panel shows information about the k-Means operator.

```

graph LR
    RD[Retrieve Data_Tugas...]
    RD --> Clustering[Clustering]
    Clustering --> SVD[SVD]
  
```

- Nilai Eigenvalue

The screenshot shows the RapidMiner Studio interface with the process titled "Tugas\_NilaiUjian - Excel". The "SVD (SVD)" tab is active. The "Eigenvalues" tab is selected, showing the following data:

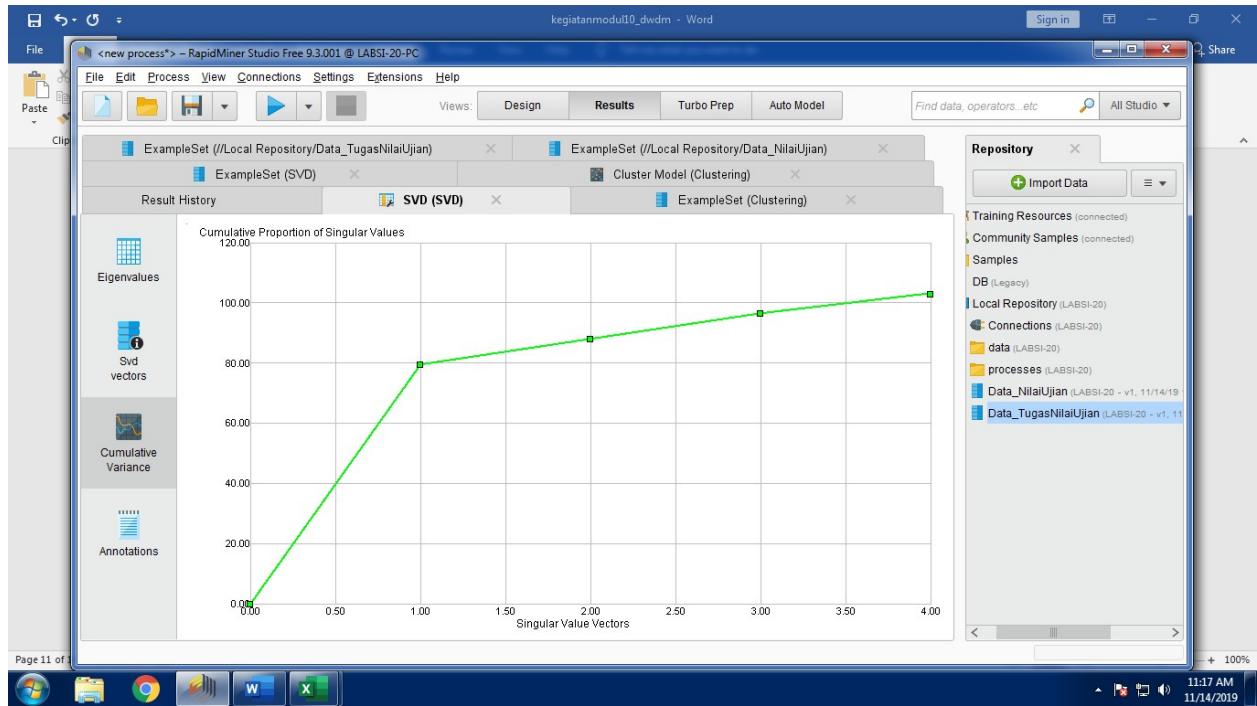
Component	Singular Value	Proportion of Singular V...	Cumulative Singular Val...	Cumulative Proportion o...
SVD 1	79.590	0.771	79.590	0.771
SVD 2	8.517	0.083	88.107	0.854
SVD 3	8.463	0.082	96.570	0.936
SVD 4	6.593	0.064	103.163	1.000

- Nilai Svd Vectors

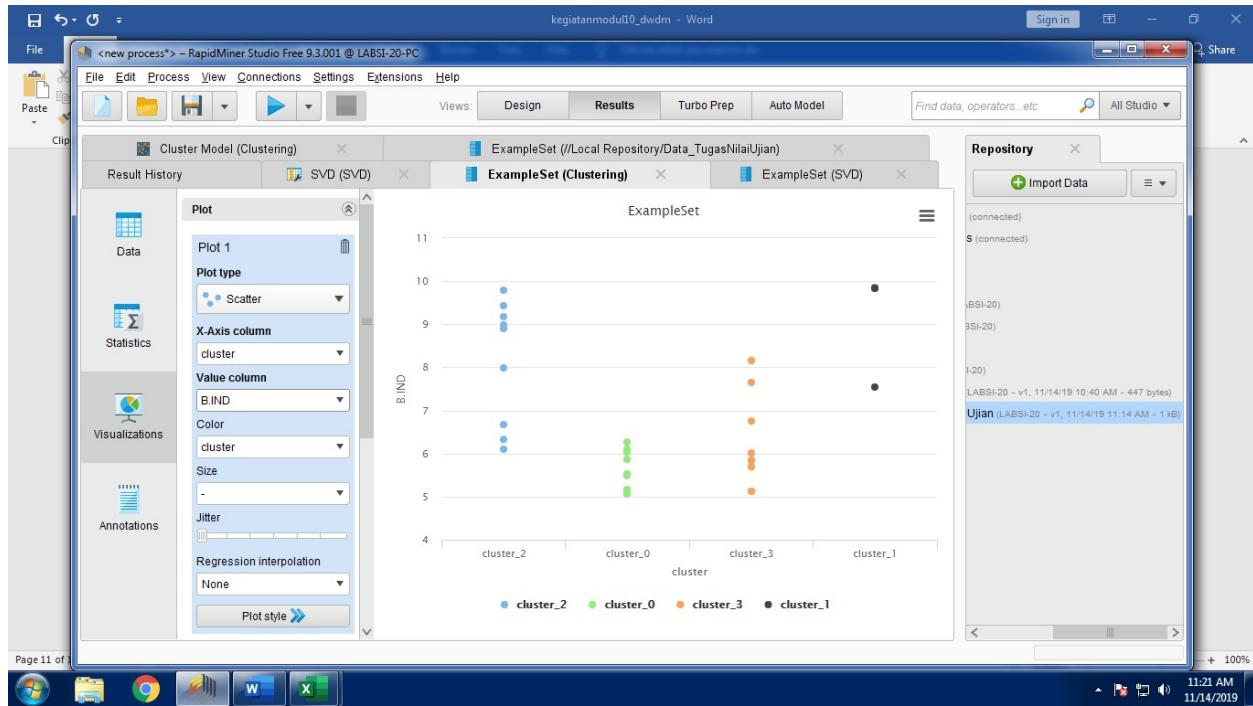
The screenshot shows the RapidMiner Studio interface with the process titled "kegiatanmodul10\_dwdm - Word". The "SVD (SVD)" tab is active. The "Svd vectors" tab is selected, showing the following data:

Attribute	SVD Vector 1	SVD Vector 2	SVD Vector 3
B.IND	0.481	-0.218	-0.699
B.ING	0.483	0.764	-0.175
MTK	0.495	-0.604	0.104
IPA	0.539	0.064	0.685

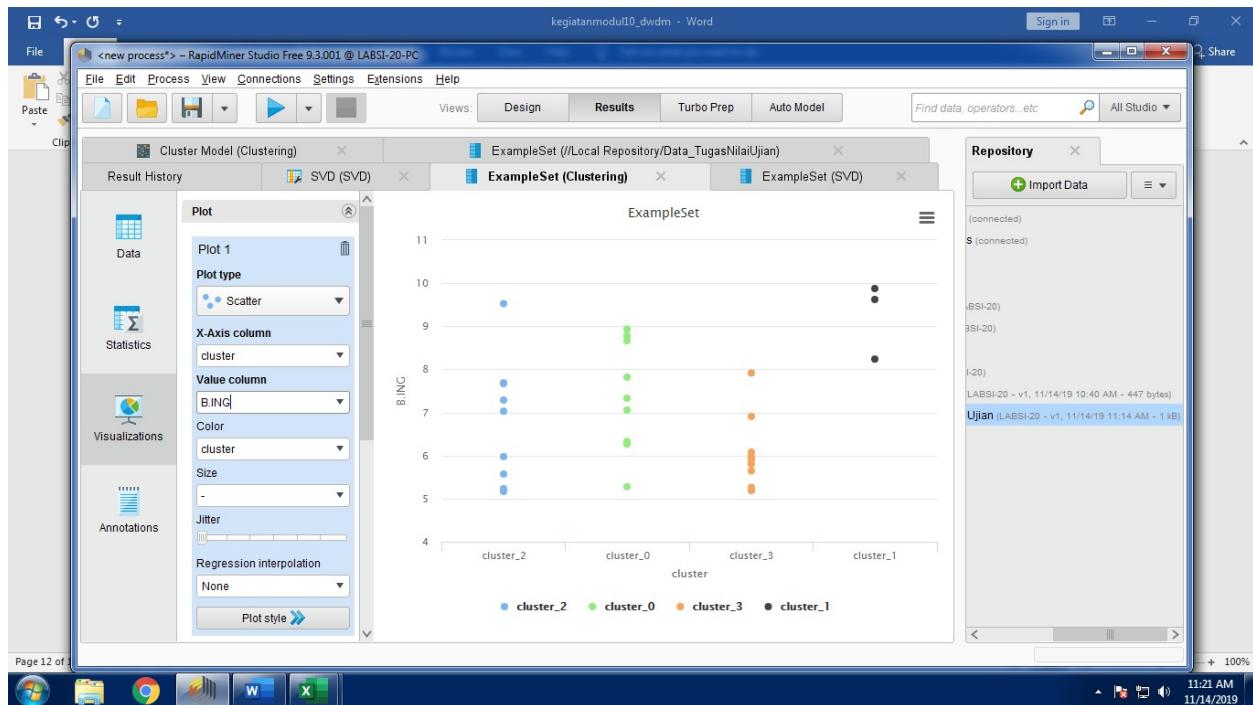
- Nilai Cumulative Variance



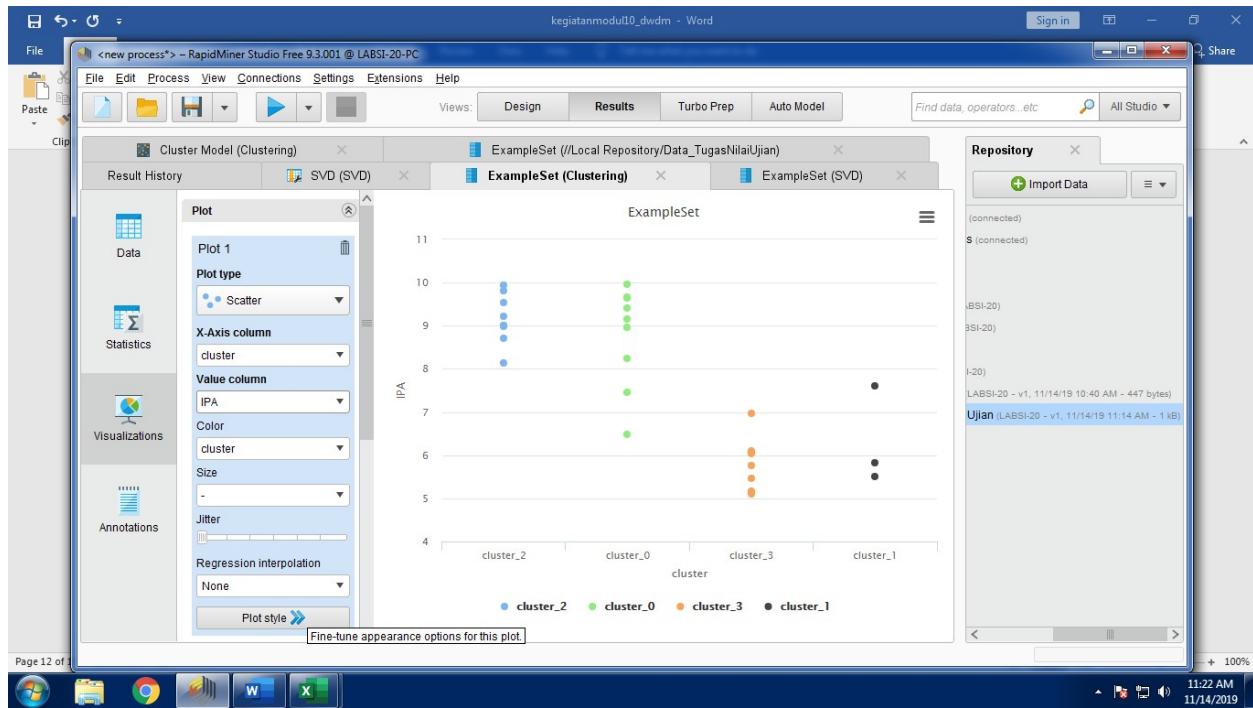
- Kelompok siswa bidang B.INDO



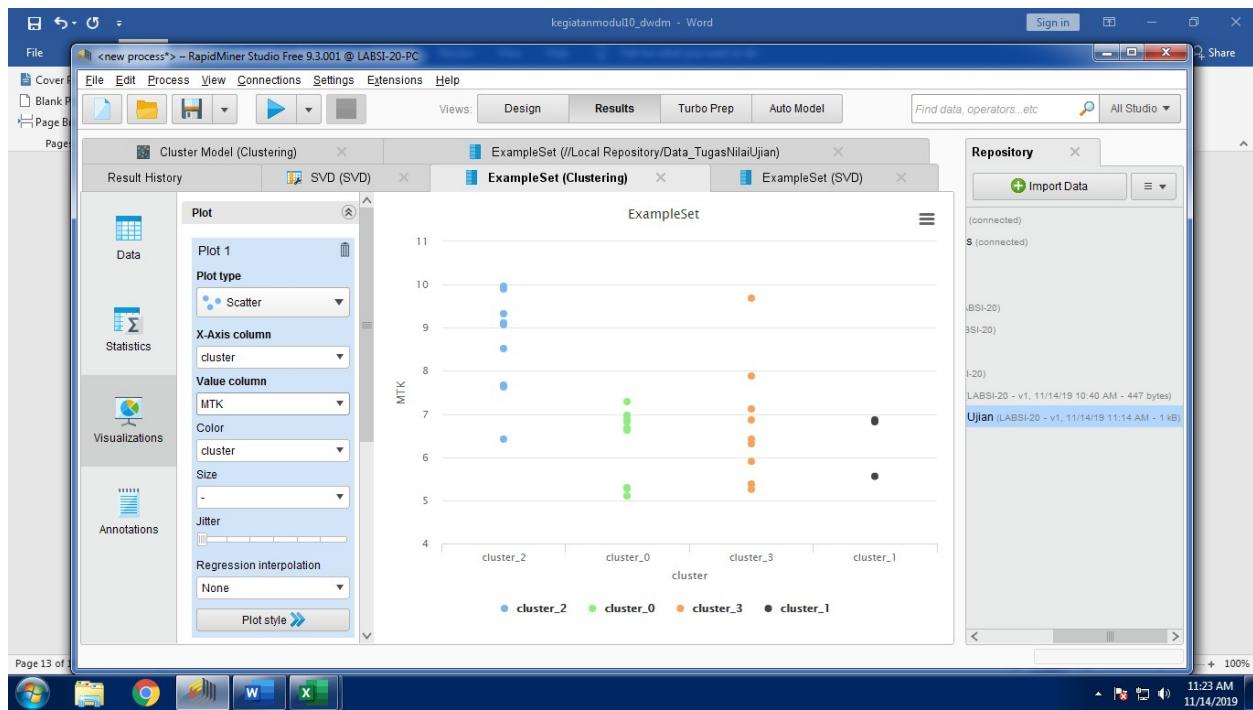
- Kelompok siswa kelompok B.ING



- Kelompok siswa bidang IPA



- Kelompok siswa bidang MTK



- Masing-masing nama siswa yang terdapat dalam kelompok cluster 0, cluster 1, cluster 2, cluster 3.

Screenshot of the RapidMiner Studio Free 9.3.001 interface showing the results of a clustering process. The main window displays a table of student data with their names, cluster assignments, and grades. The Repository panel on the right shows the imported dataset.

**Repository**

- (connected)
- S (connected)
- LABSI-20
- BSI-20
- I-20
- Ujian (LABSI-20 - v1, 11/14/19 10:40 AM - 447 bytes)

Row No.	NAMA	cluster ↑	BJND	B.ING	MTK	IPA
2	AGUS	cluster_0	5.869	5.275	6.666	8.243
5	WATI	cluster_0	6.039	7.820	5.111	9.961
7	EKO	cluster_0	5.485	8.767	5.277	9.657
14	DANI	cluster_0	5.170	7.338	6.863	9.401
15	AHMAD	cluster_0	5.140	8.933	6.625	6.490
16	BAYU	cluster_0	5.062	6.268	7.287	8.958
23	SARAH	cluster_0	6.105	7.050	6.814	7.449
25	BAMBANG	cluster_0	5.523	6.326	5.300	9.633
26	HADI	cluster_0	6.266	8.670	6.977	9.151
17	RISA	cluster_1	9.846	9.611	6.862	7.611
19	YANI	cluster_1	7.547	9.881	6.825	5.510
22	JONO	cluster_1	9.832	8.238	5.557	5.830

ExampleSet (30 examples, 2 special attributes, 4 regular attributes)

RapidMiner Studio Free 9.3.001 @ LABSI-20-PC

kegiatanmodul10\_dwdm - Word

File Edit Process View Connections Settings Extensions Help

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

Cluster Model (Clustering) ExampleSet (/Local Repository/Data\_TugasNilaiUjian) ExampleSet (Clustering) ExampleSet (SVD)

Result History SVD (SVD) ExampleSet (SVD)

Repository Import Data

Data

Open in Turbo Prep Auto Model Filter (30 / 30 examples): all

Row No.	NAMA	cluster ↑	B.IND	B.ING	MTK	IPA
22	JONO	cluster_1	9.832	8.238	5.557	5.830
1	JOKO	cluster_2	6.330	7.029	9.950	9.001
4	DYAH	cluster_2	6.671	5.570	9.109	8.714
8	YANTO	cluster_2	9.438	9.526	7.819	9.803
10	MAHMUD	cluster_2	9.182	5.976	6.424	8.129
12	SANTI	cluster_2	8.975	7.280	8.514	9.545
13	DIAN	cluster_2	6.098	7.660	9.331	9.226
20	RATIH	cluster_2	9.785	5.239	9.893	8.987
21	INDAH	cluster_2	8.901	5.171	9.076	9.005
27	NANA	cluster_2	7.979	7.688	7.679	9.945
3	SUSI	cluster_3	5.692	7.916	5.903	5.166
6	IKA	cluster_3	7.658	6.069	6.309	6.972

ExampleSet (30 examples, 2 special attributes, 4 regular attributes)

Page 14 of 14 11:24 AM 11/14/2019

The screenshot shows the RapidMiner Studio Free application window. The main area displays a clustered dataset with 30 examples across four regular attributes (B.IND, B.ING, MTK, IPA) and two special attributes (cluster). The data is grouped into three clusters: cluster\_1 (22 rows), cluster\_2 (14 rows), and cluster\_3 (4 rows). The 'Results' tab is selected. On the left, there are tabs for 'Data', 'Statistics', 'Visualizations', and 'Annotations'. The 'Repository' pane on the right shows a connection to 'LABSI-20' and 'Ujian'. The bottom status bar indicates the page is 14 of 14, the time is 11:24 AM, and the date is 11/14/2019.

The screenshot shows the RapidMiner Studio interface. The main window displays a data table titled "ExampleSet (//Local Repository/Data\_TugasNilaiUjian)". The table has columns: Row No., NAMA, cluster ↑, B.IND, B.ING, MTK, and IPA. The data shows 30 rows of student names and their scores, grouped into three clusters (cluster\_2 and cluster\_3). The "Annotations" tab is selected in the sidebar.

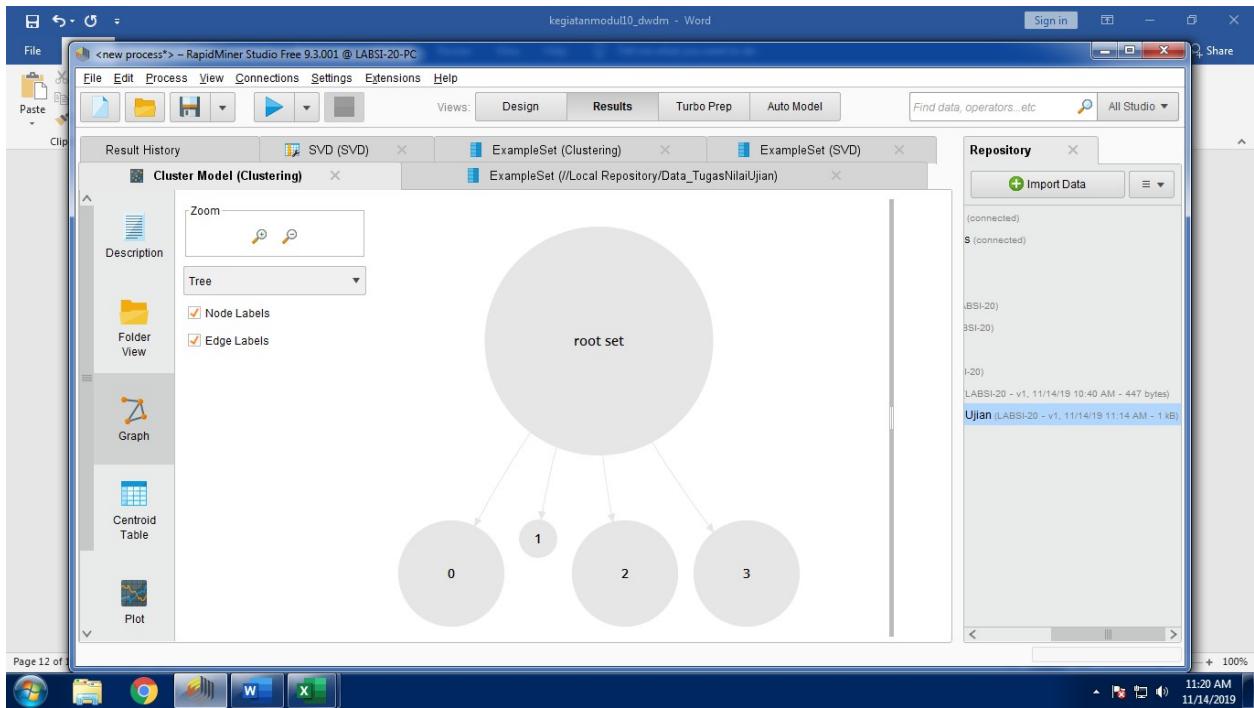
Row No.	NAMA	cluster ↑	B.IND	B.ING	MTK	IPA
20	RATIH	cluster_2	9.785	5.239	9.893	8.987
21	INDAH	cluster_2	8.901	5.171	9.076	9.005
27	NANA	cluster_2	7.979	7.688	7.679	9.945
3	SUSI	cluster_3	5.692	7.916	5.903	5.166
6	IKA	cluster_3	7.658	6.069	6.309	6.972
9	WAWAN	cluster_3	5.839	5.907	5.256	5.454
11	BUDI	cluster_3	8.155	5.177	6.863	5.152
18	RANI	cluster_3	5.126	6.906	9.694	6.069
24	RAMA	cluster_3	5.127	5.262	6.411	5.104
28	FEBRI	cluster_3	6.754	5.798	7.874	5.763
29	DENI	cluster_3	5.846	5.969	7.120	6.037
30	TONI	cluster_3	6.011	5.644	5.378	6.099

ExampleSet (30 examples, 2 special attributes, 4 regular attributes)

- Description

The screenshot shows the RapidMiner Studio interface. The main window displays the "Cluster Model (Clustering)" description. It states: "Cluster 0: 9 items", "Cluster 1: 3 items", "Cluster 2: 9 items", "Cluster 3: 9 items", and "Total number of items: 30". The "Description" tab is selected in the sidebar.

- Graph

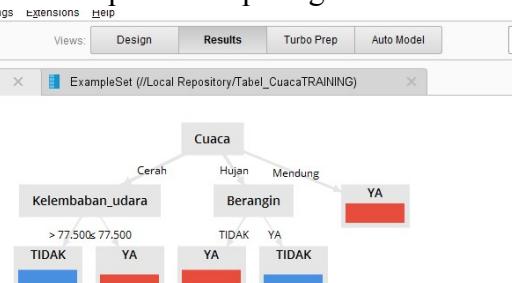


## H. MODUL 11

### A. Kegiatan praktikum

#### 1. Induksi Aturan Data Cuaca

- Menggunakan proses dari praktikum modul 9 dan menghasilkan sebuah pohon keputusan seperti gambar dibawah.



- b. Mengubah parameter sehingga menghasilkan sebuah induksi aturan dari data training yang disebut sebagai Rule Model

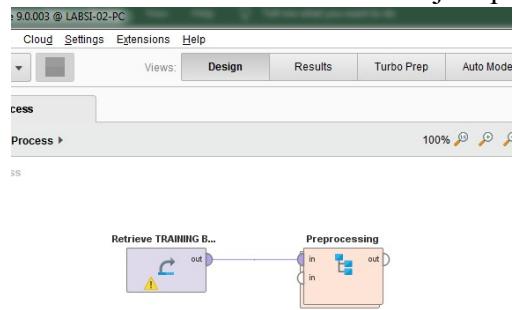
```

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.
  
```

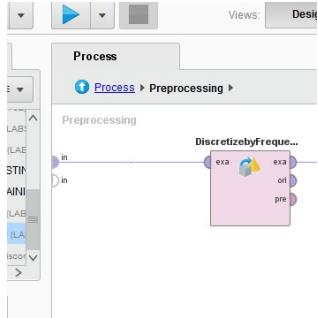
## 2. Aturan Asosiasi Data Cuaca

- a. Menggunakan DataCuaca\_Training dan menggunakan parameter subprocess ke dalam area. Lalu ubah nama menjadi preprocessing.



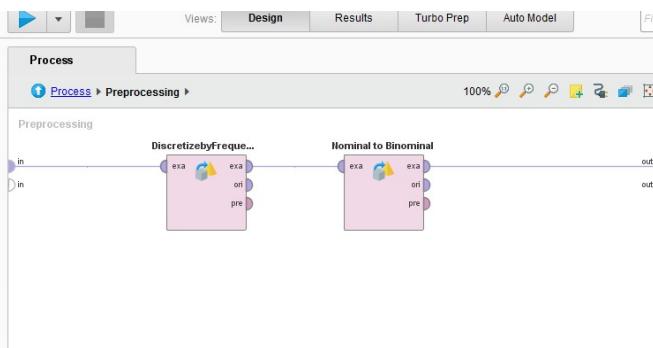
- b. Klik ganda pada operator preprocesssing sehingga masuk pada nested chain. Lalu rename

menjadi “DiscretizebyFrequency” dan biarkan nilai parameter number of bins = 2.

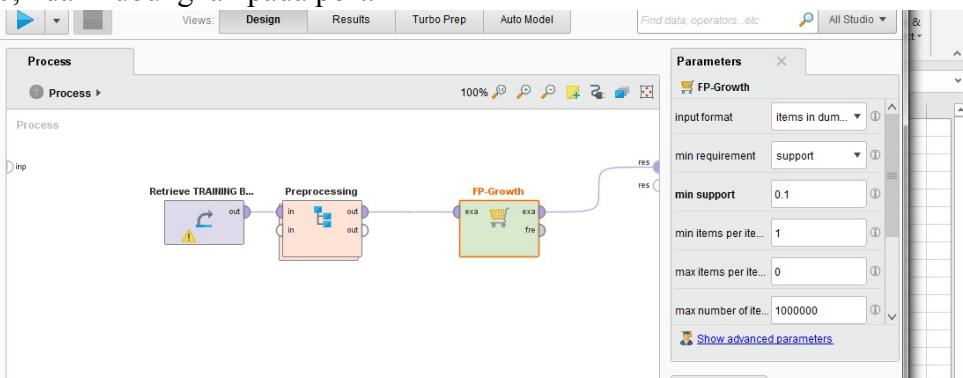


- c. Mengubah nama operator menjadi “Nominal2Binomial” dan menghubungkan operator

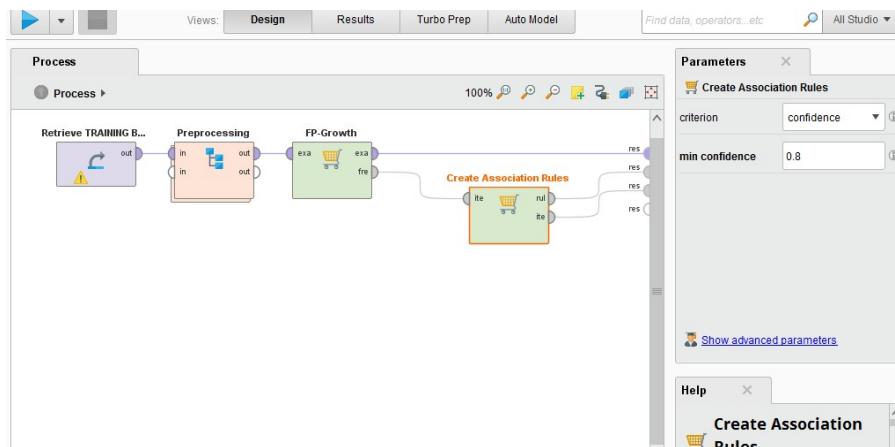
sebelumnya dengan masukan examination pada operator ini dengan panel out.



- d. Tambahkan parameter Fp-Growth dan ubah parameter pada min support = 0,1 dan hubungkan pada port.



- e. Tambahkan parameter Create Association Rules dan menghubungkan operator-operator tersebut.



f. Dapat dilihat hasil-hasil aturan asosiasi sebagai berikut.

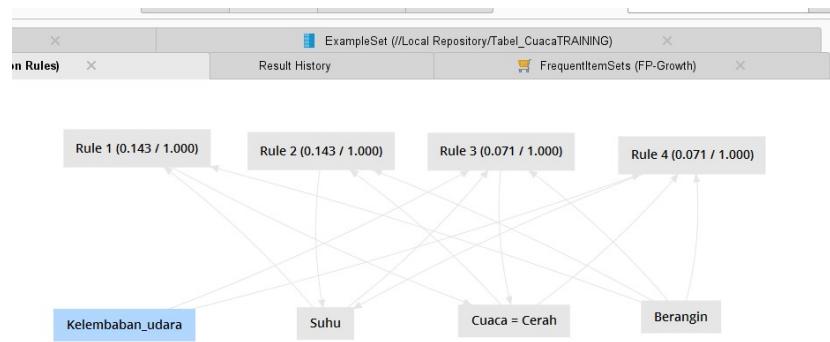
No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain
1	Berangin, Suhu	Cuaca = Cerah	0.143	1	1	-0.143
2	Berangin, Cuaca = Cerah	Suhu	0.143	1	1	-0.143
3	Kelembaban_udara, Berangin, Suhu	Cuaca = Cerah	0.071	1	1	-0.071
4	Kelembaban_udara, Berangin, Cuaca = Cerah	Suhu	0.071	1	1	-0.071

Dapat dilihat bahwa jumlah aturan asosiasi yang terbentuk adalah 23 set dan jumlah total maximal size = 4 yang terdiri dari 4 buah itemset.

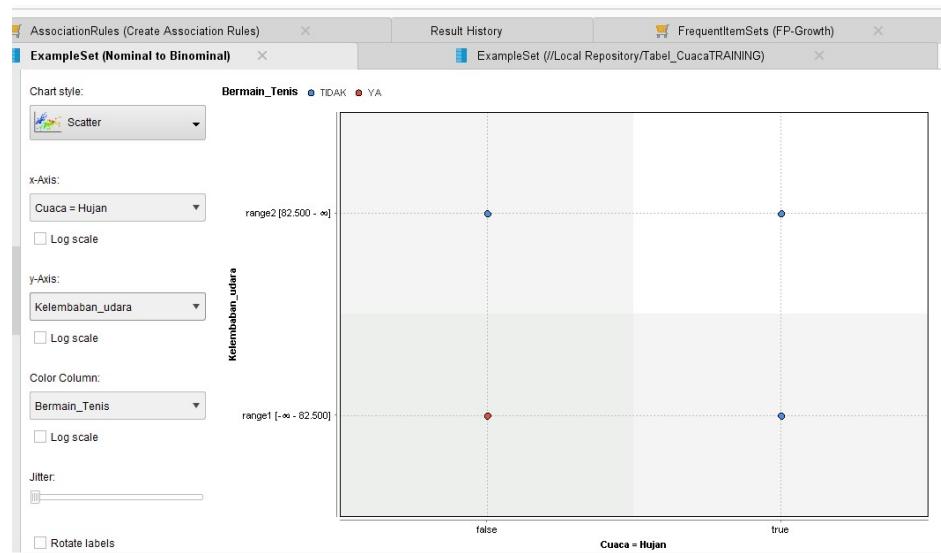
#### - Table views

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

#### - Graph views

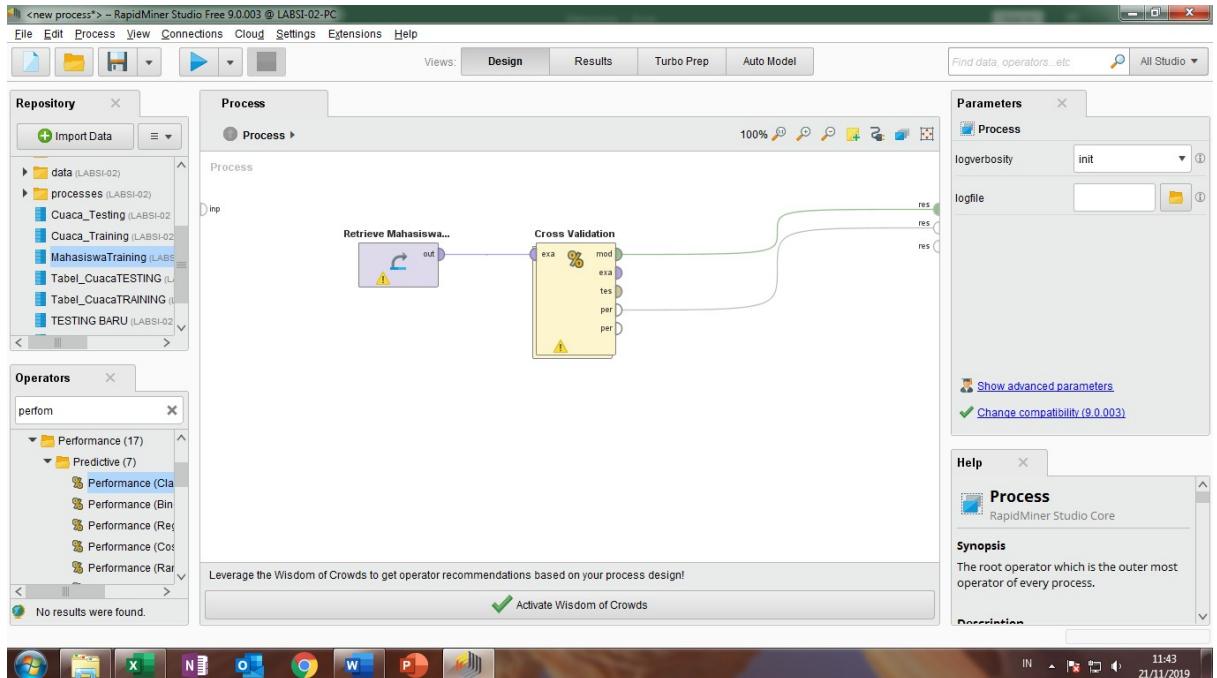


## - ExampleSet

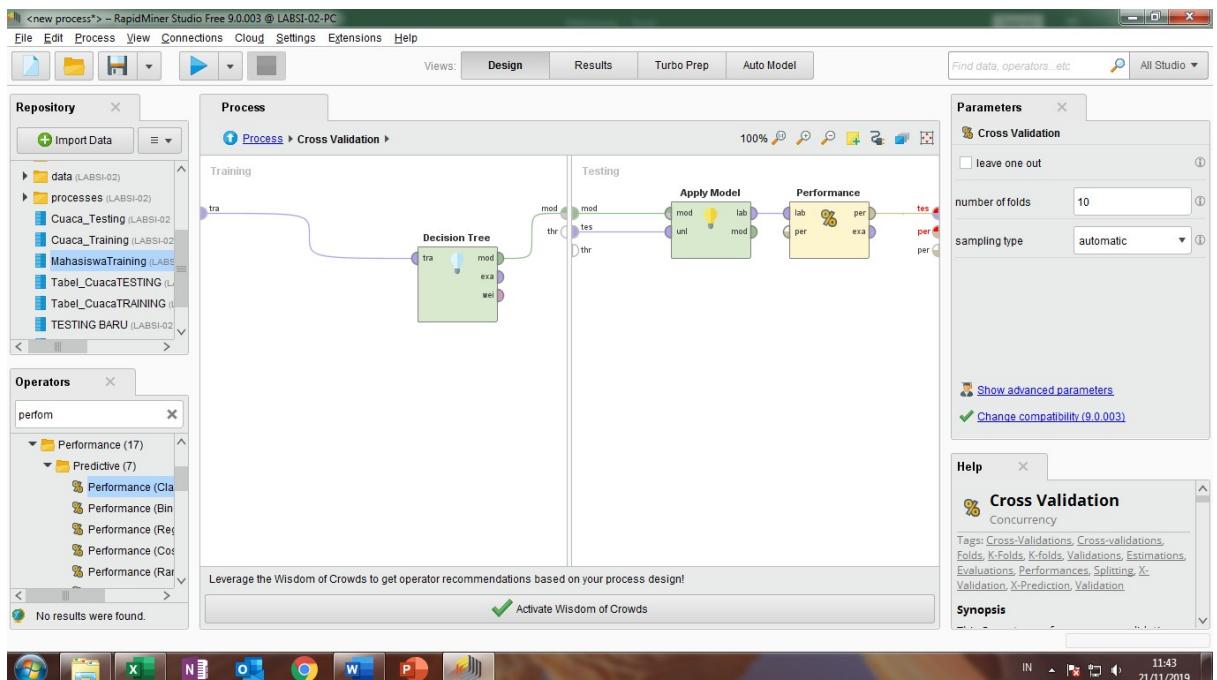


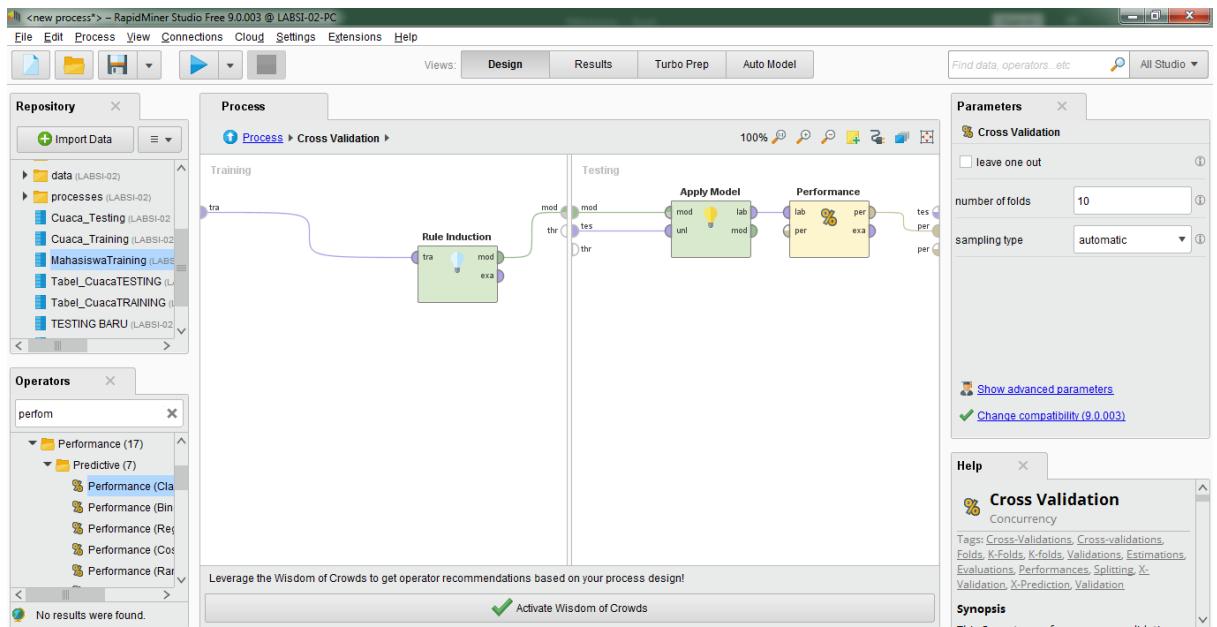
## B. Tugas Praktikum

- Membuat operator dengan menggunakan decision tree.



- Lalu klik ganda pada decision tree dan klik ganda pada operator tersebut.





**RapidMiner Studio Free 9.0.003 @ LABSI-02-PC**

**Result History**

ExampleSet (//Local Repository/Tabel\_CuacaTRAINING) | PerformanceVector (Performance) | RuleModel (Rule Induction)

Criterion: accuracy

Table View | Plot View

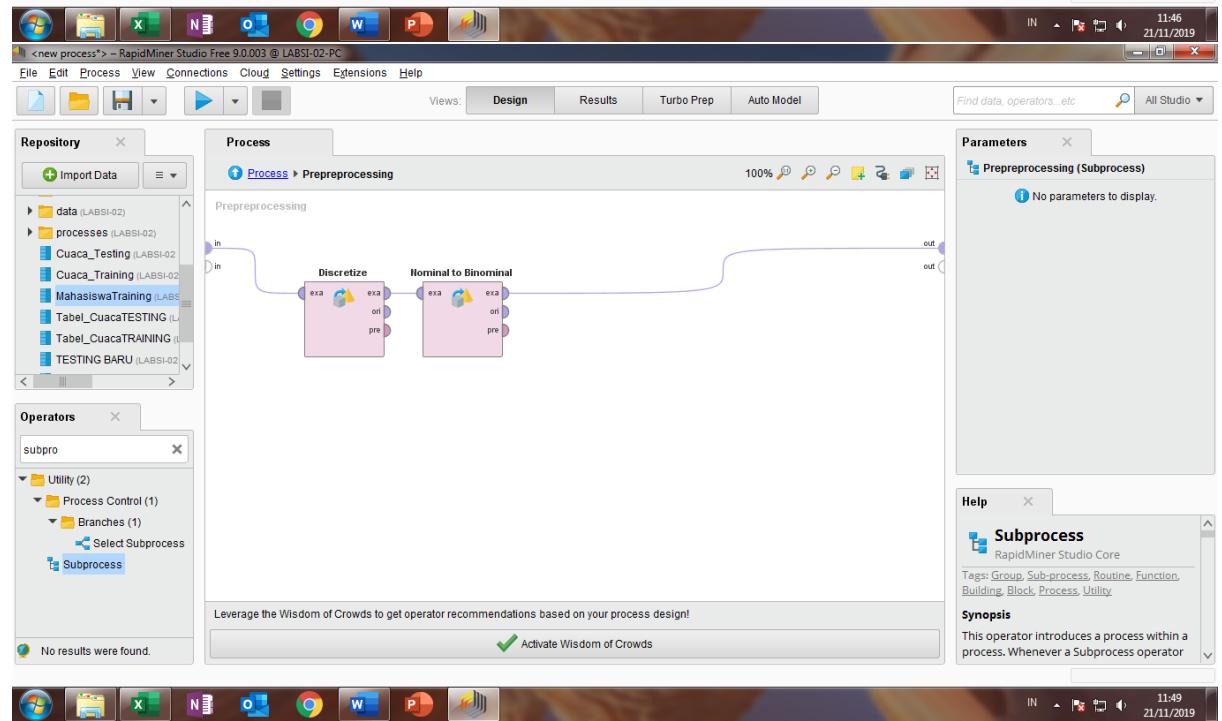
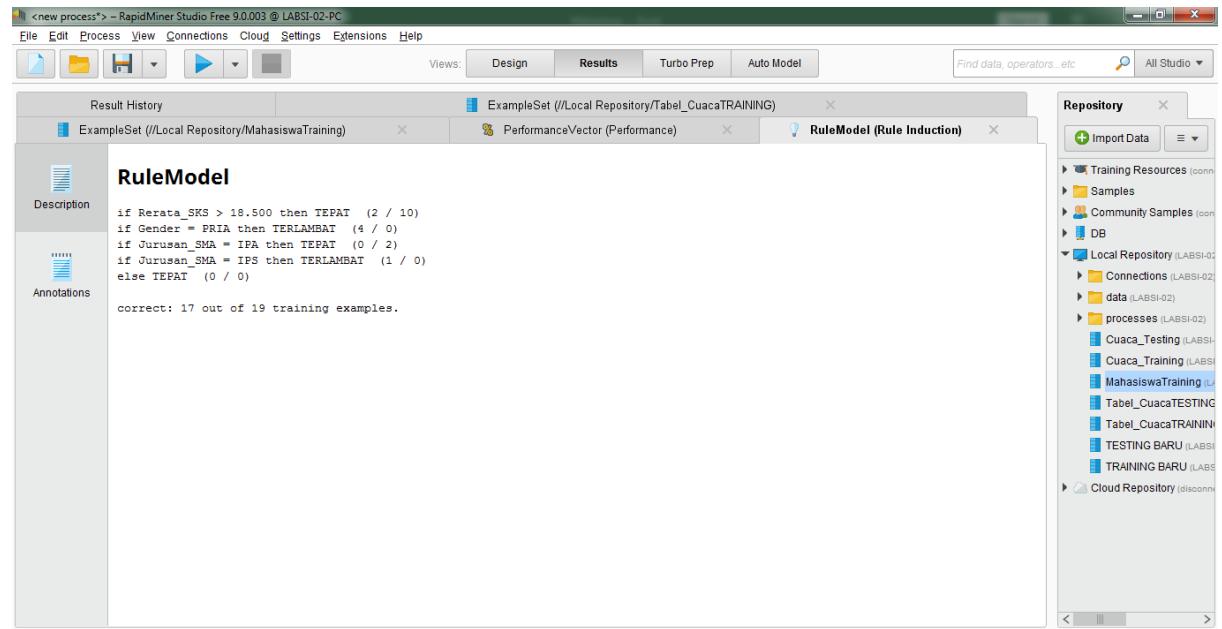
accuracy: 65.00% +/- 32.02% (micro average: 65.00%)

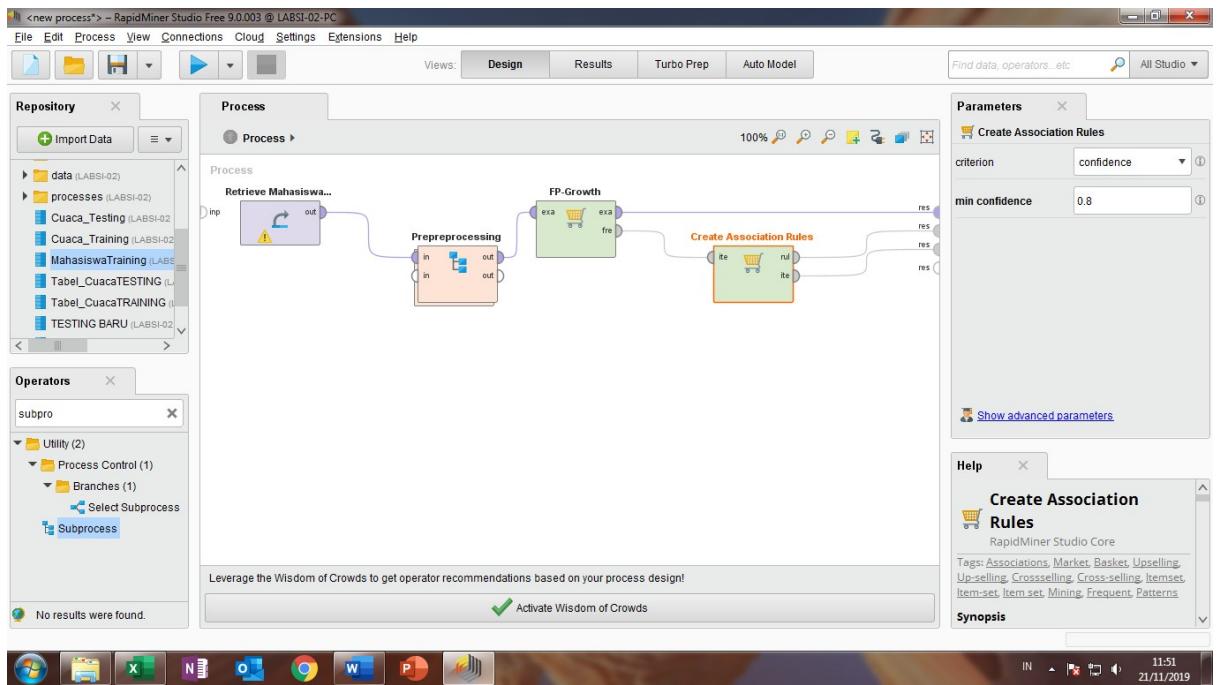
	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	4	4	50.00%
pred. TEPAT	3	9	75.00%
class recall	57.14%	69.23%	

**Repository**

- Training Resources (conn)
- Samples
- Community Samples (conn)
- DB
- Local Repository (LABSI-02)
  - Connections (LABSI-02)
  - data (LABSI-02)
  - processes (LABSI-02)
  - Cuaca\_Testing (LABSI-02)
  - Cuaca\_Training (LABSI-02)
  - MahasiswaTraining (LABSI-02)
  - Tabel\_CuacaTESTING (LABSI-02)
  - Tabel\_CuacaTRAINING (LABSI-02)
  - TESTING BARU (LABSI-02)
  - TRAINING BARU (LABSI-02)
- Cloud Repository (disconn)

## - Rule Model





## - Table Views

The screenshot shows the RapidMiner Studio Free interface with two tables side-by-side:

**FrequentItemSets (FP-Growth) View:**

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

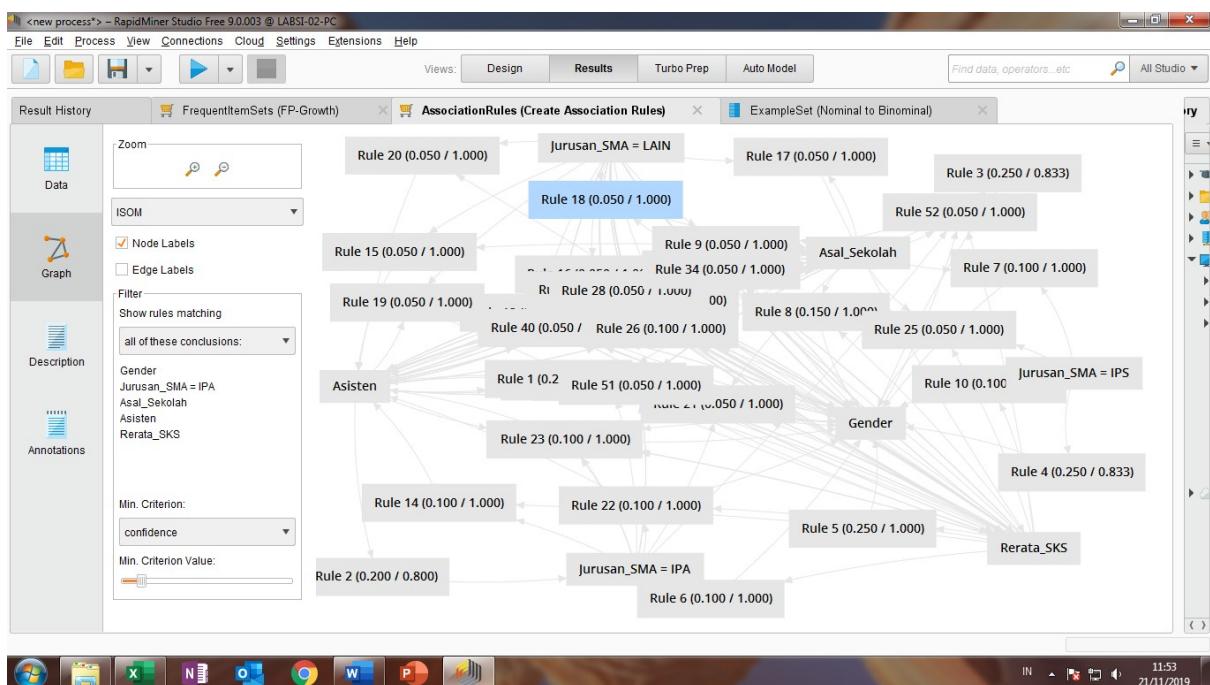
**AssociationRules (Create Association Rules) View:**



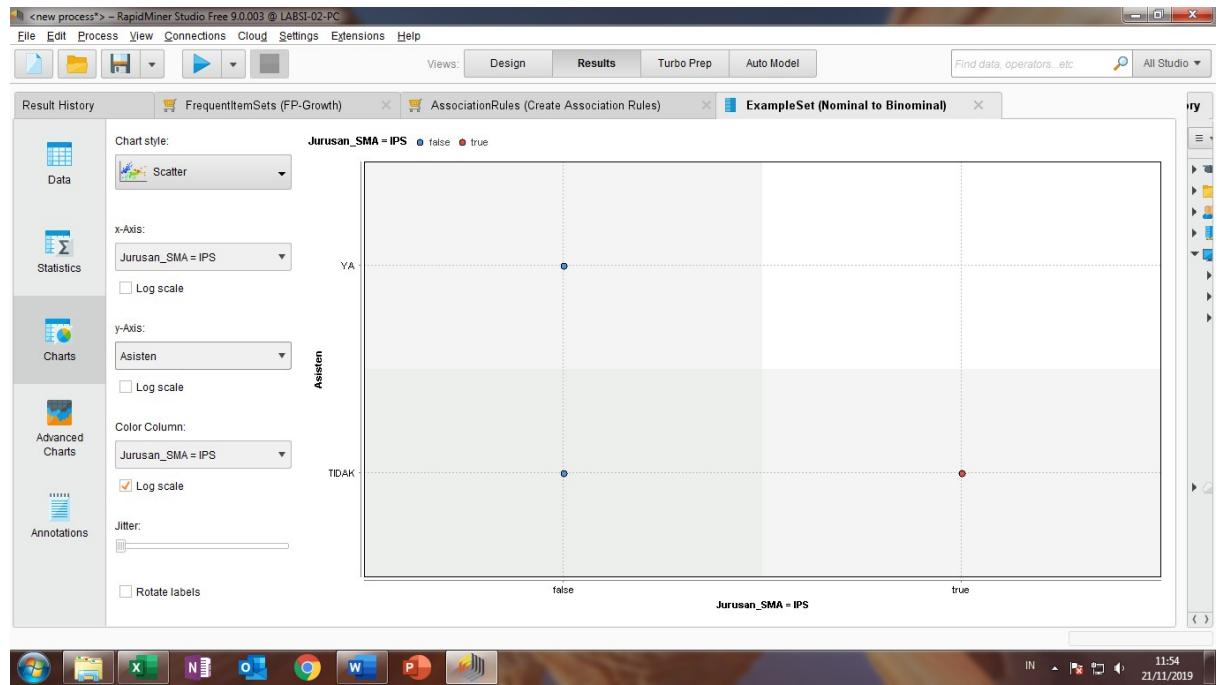
The "Repository" panel on the right lists various datasets and processes available in the workspace.

FrequentItemSets (FP-Growth)							AssociationRules (Create Association Rules)									
Result History		Views:			Design		Results		Turbo Prep		Auto Model		Find data, operators...etc		All Studio	
Data	Show rules matching all of these conclusions:	No.	Premises	Conclusion			Support		Confidence		LaPlace		Gain			
Gender	Jurusan_SMA = IPA	3	Asal_Sekolah	Gender			0.250		0.833		0.962		-0.350			
Asisten	Asal_Sekolah	4	Jurusan_SMA = IPS	Gender			0.250		0.833		0.962		-0.350			
Rerata_SKS	Rerata_SKS	5		Gender			0.250		1		1		-0.250			
	Jurusan_SMA = IPA, Rerata_SKS	6		Gender			0.100		1		1		-0.100			
	Asal_Sekolah, Jurusan_SMA = IPS	7		Gender			0.100		1		1		-0.100			
	Asal_Sekolah, Rerata_SKS	8		Gender			0.150		1		1		-0.150			
	Asal_Sekolah, Jurusan_SMA = LAIN	9		Gender			0.050		1		1		-0.050			
	Jurusan_SMA = IPS, Rerata_SKS	10		Gender			0.100		1		1		-0.100			
	Asisten, Rerata_SKS	11		Gender			0.150		1		1		-0.150			
	Asisten, Jurusan_SMA = LAIN	12		Gender			0.050		1		1		-0.050			
	Rerata_SKS, Jurusan_SMA = LAIN	13		Gender			0.050		1		1		-0.050			
	Jurusan_SMA = IPA, Rerata_SKS	14		Asisten			0.100		1		1		-0.100			
	Asal_Sekolah, Jurusan_SMA = LAIN	15		Asisten			0.050		1		1		-0.050			
	Asisten, Jurusan_SMA = LAIN	16		Asal_Sekolah			0.050		1		1		-0.050			
	Asal_Sekolah, Jurusan_SMA = LAIN	17		Rerata_SKS			0.050		1		1		-0.050			
	Rerata_SKS, Jurusan_SMA = LAIN	18		Asal_Sekolah			0.050		1		1		-0.050			

### - Graph View



## - Example Set



## I. MODUL 12

### Kegiatan Praktikum

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

Import Data - Select the cells to import.



### Select the cells to import.

Sheet: Sheet1 ▾

Cell range: A:D

Select All

Define header row:

1 ▾

	A	B	C	D
1	NO_SISWA	NAMA	LAMA BELAJAR(JAM)	NILAI
2	S-101	JOKO	15.000	783.000
3	S-102	AGUS	18.000	877.000
4	S-103	SUSI	7.000	505.000
5	S-104	DYAH	9.000	860.000
6	S-105	WATI	15.000	968.000
7	S-106	IKA	17.000	793.000
8	S-107	EKO	10.000	752.000
9	S-108	YANTO	5.000	571.000
10	S-109	WAWAN	8.000	667.000
11	S-110	MAHMUD	15.000	723.000

Previous

Next

Cancel

Import Data - Format your columns.

### Format your columns.

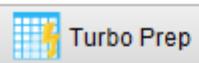
Replace errors with missing values  ⓘ

NO_SISWA <small>polynomial id</small>	NAMA <small>polynomial</small>	LAMA BELAJAR(JAM) <small>integer</small>	NILAI <small>integer label</small>
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

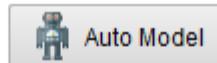
 no problems.

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

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

**Process**

Process

Process

Retrieve Data\_Lama...

Linear Regression

Parameters

Linear Regression

min tolerance 0.05

ridge 1.0E-8

Show advanced parameters

Result History

LinearRegression (Linear Regression)

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	****

## LinearRegression

21.608 \* 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

Import Data - Select the cells to import.

Select the cells to import.

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

	A	B	C
1	NO_SISWA	NAMA	LAMA BELAJAR(JAM)
2	S-111	BUDI	12.000
3	S-112	SANTI	13.000
4	S-113	DIAN	14.000
5	S-114	DANI	11.000
6	S-115	AHMAD	5.000
7	S-116	BAYU	13.000
8	S-117	RISA	9.000
9	S-118	RANI	10.000
10	S-119	YANI	10.000
11	S-120	RATIH	9.000

← Previous → Next ✖ Cancel

Import Data - Format your columns.

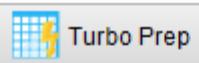
Format your columns.

Replace errors with missing values ⓘ

NO_SISWA polynomial id	NAMA polynomial	LAMA BELAJAR(JAM) integer
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

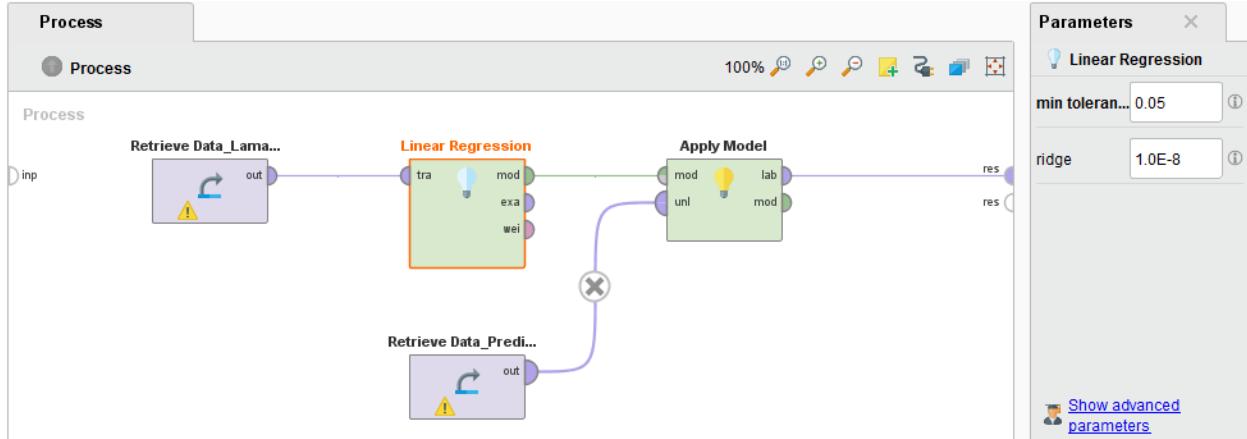


Turbo Prep



Auto Model

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

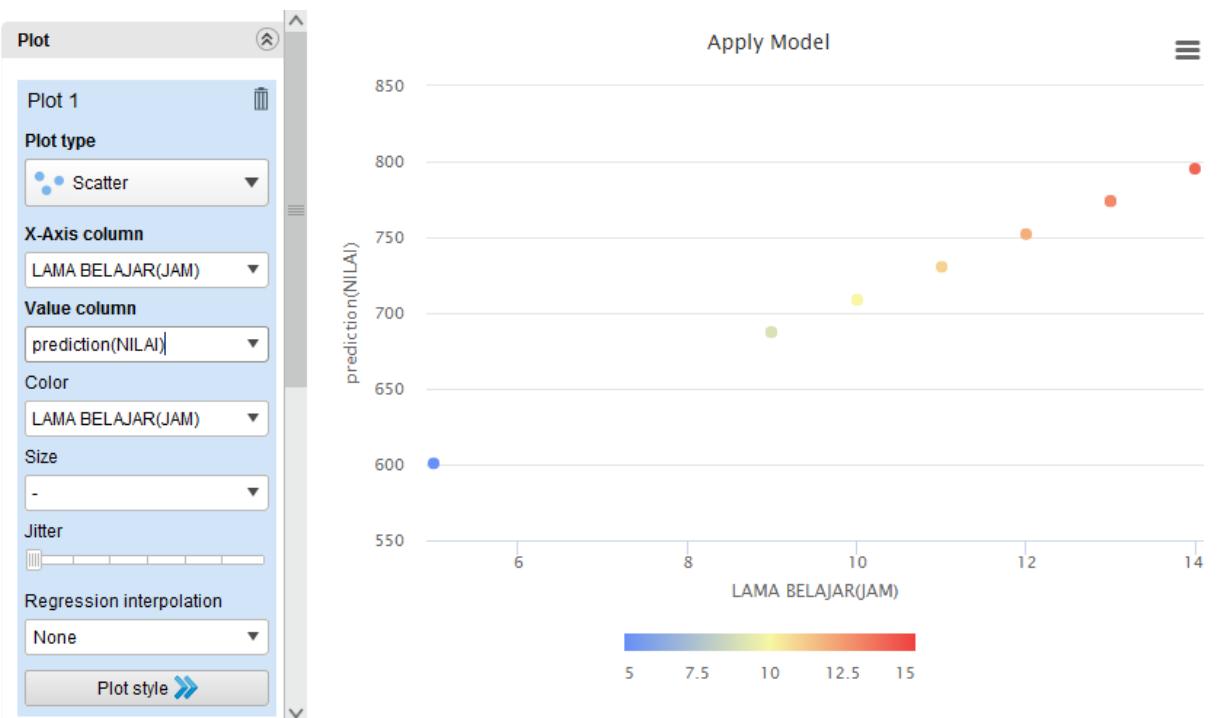


Open in

Turbo Prep

Auto Model

Row No.	NO_SISWA	prediction(NILAI)	LAMA BELAJAR(JAM)
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



A	B	C	D	E
NO_SISWA	NAMA	LAMA BELAJAR(JAM)	Tabel	Model Regression
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

## TUGAS

Import Data - Format your columns.

**Format your columns.**

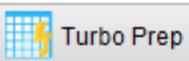
Replace errors with missing values (i)

NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	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.

Previous Next ✗ Cancel

Open in



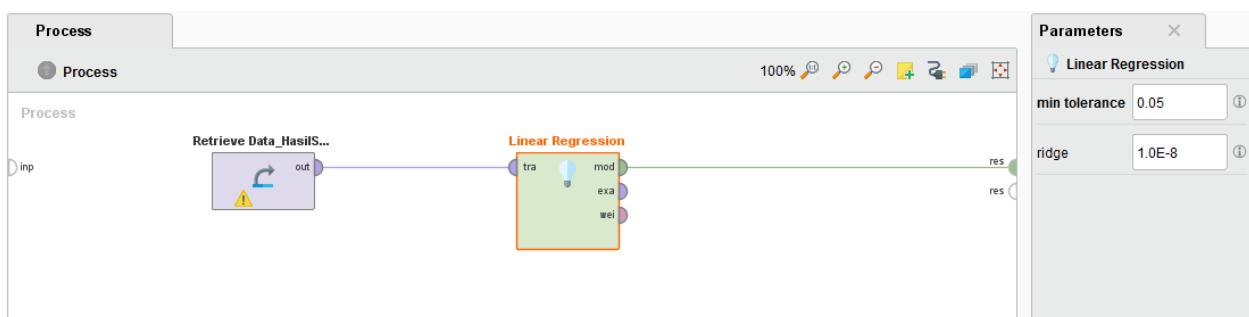
Turbo Prep



Auto Model

Row No.	NO RESPON...	DAYA BELI(R...	PENDAPATA...	JUMLAH AN...
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
14	14	1125000	1500000	4
15	15	1115000	1300000	7

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



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

## Linear Regression

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

NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	DAYA BELI (RUPIAH)	Y
1	1000000	6	834000	845623.3
2	1400000	7	1200000	1189031
3	200000	3	134000	111000.4
4	1400000	6	1167000	1141223
5	500000	3	334000	332700.4
6	1700000	5	1360000	1315116
7	400000	3	267000	258800.4
8	1900000	5	1520000	1462916
9	300000	3	200000	184900.4
10	500000	4	375000	380508
11	700000	7	600000	671730.9
12	1900000	3	1267000	1367300
13	800000	4	600000	602208
14	1500000	4	1125000	1119508
15	1300000	7	1115000	1115131

Import Data - Format your columns.

### Format your columns.

Replace errors with missing values ①

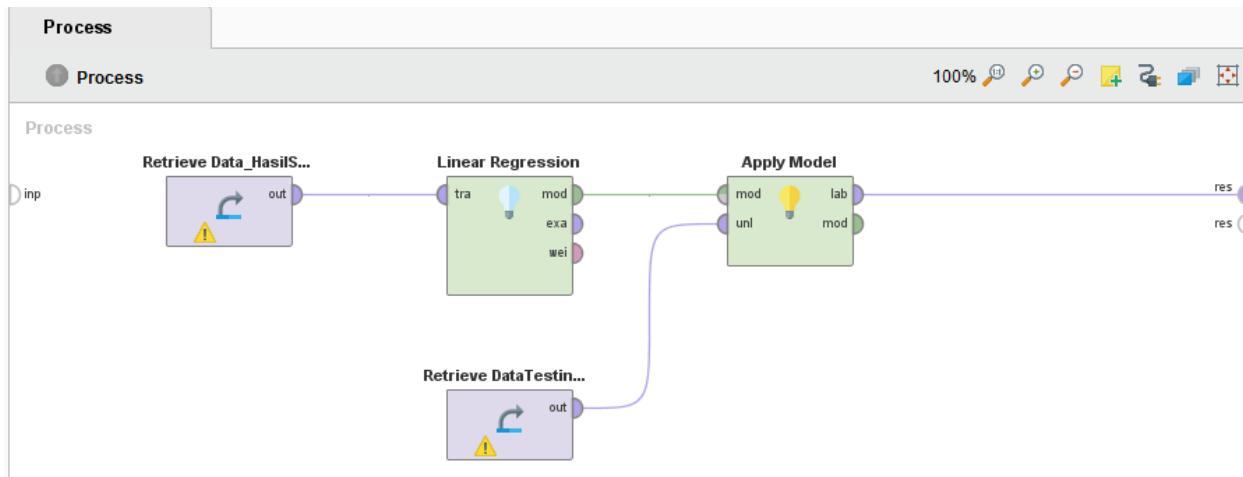
	NO RESPONDEN <small>integer id</small>	PENDAPATAN (RUPIAH) <small>integer</small>	JUMLAH ANGGOTA KELUARGA <small>integer label</small>
1	1	900000	5
2	2	800000	3
3	3	500000	2
4	4	1900000	6
5	5	600000	2
6	6	800000	5
7	7	1000000	6
8	8	1100000	4
9	9	1000000	4
10	10	500000	3

 no problems.

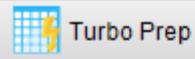
 Previous  Next  Cancel

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

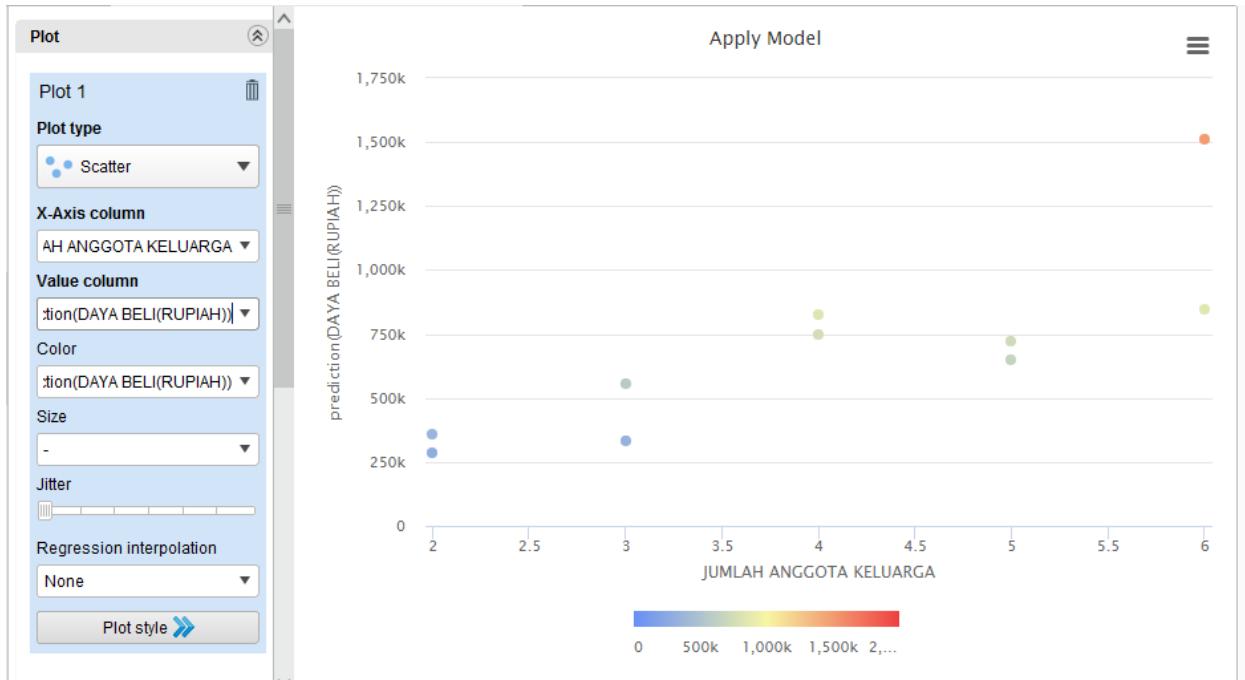


Turbo Prep



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



E2

=0.739\*B2+47807.624\*C2-180222.487

A	B	C	D	E
1	NO RESPONDEN	PENDAPATAN (RUPIAH)	JUMLAH ANGGOTA KELUARGA	TABEL
2	1	900000	5	723933.263 723915.633
3	2	800000	3	554416.056 554400.385
4	3	500000	2	284902.556 284892.761
5	4	1900000	6	1510760.476 1510723.257
6	5	600000	2	358804.515 358792.761
7	6	800000	5	650031.304 650015.633
8	7	1000000	6	845642.845 845623.257
9	8	1100000	4	823929.557 823908.009
10	9	1000000	4	750027.598 750008.009
11	10	500000	3	332710.179 332700.385
12				