

Nama : Pawitro Purbsngkoro

NIM : L200170045

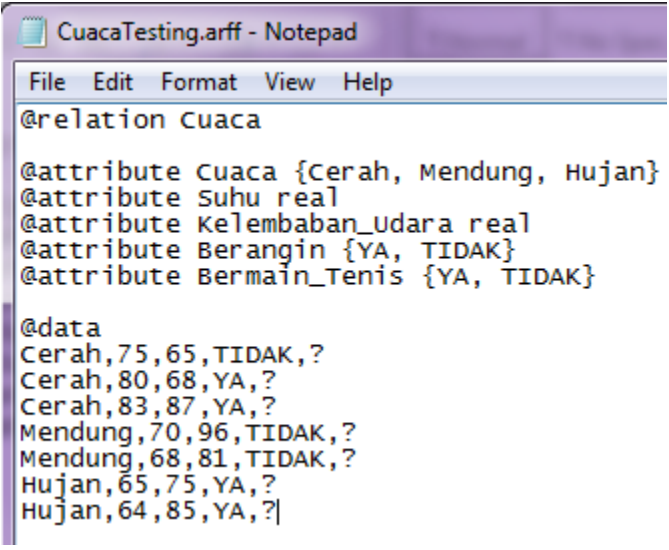
Kelas : C

MODUL08

1. 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
```

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

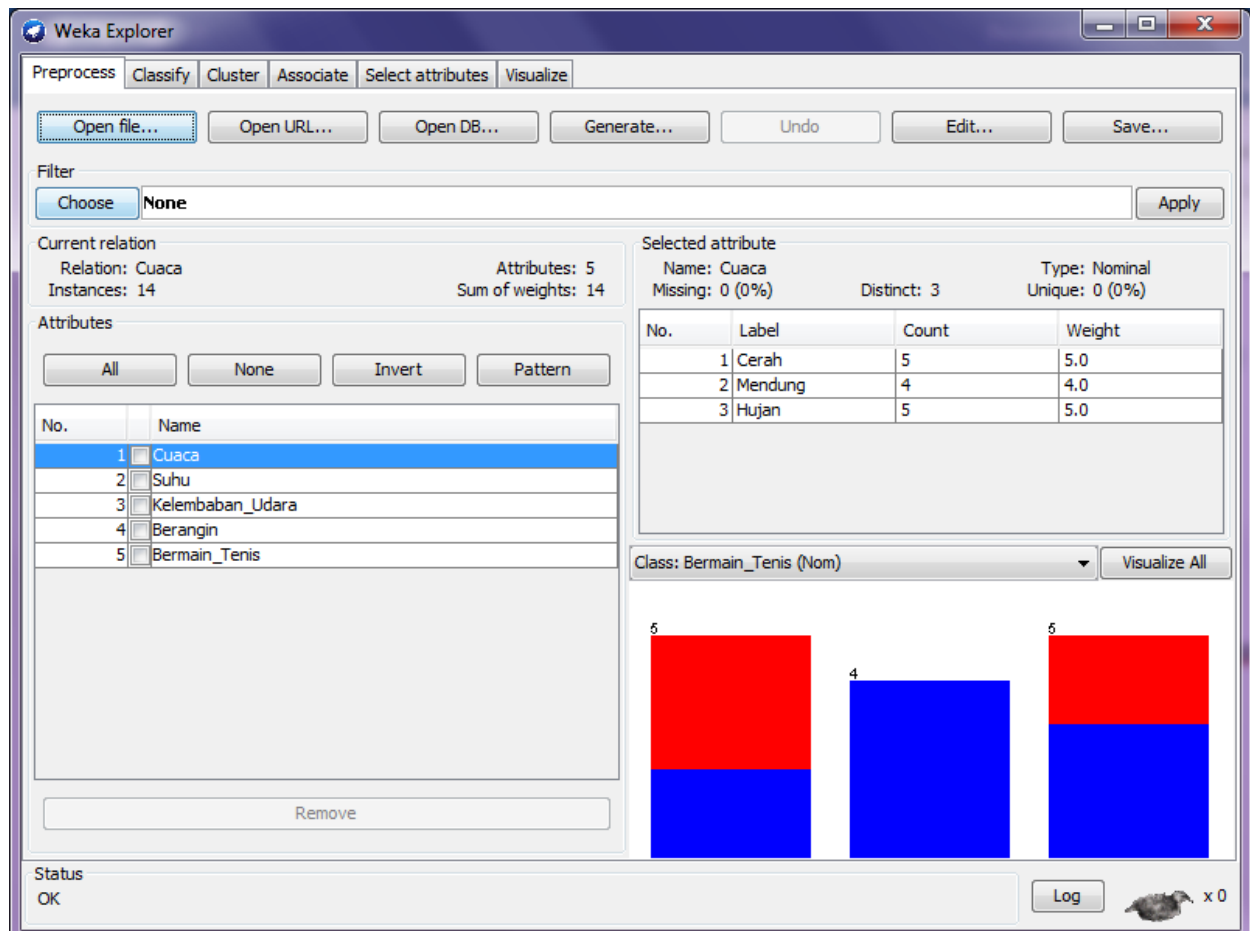


```
CuacaTesting.arff - Notepad
File Edit Format View Help
@relation Cuaca

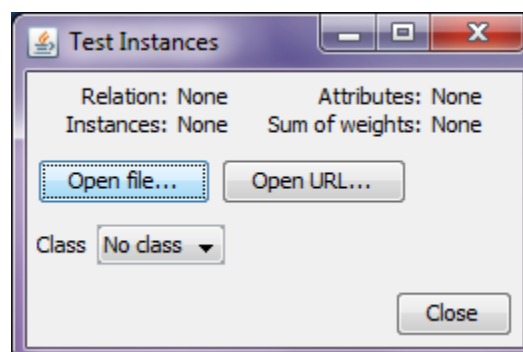
@attribute Cuaca {Cerah, Mendung, Hujan}
@attribute Suhu real
@attribute Kelembaban_Udara real
@attribute Berangin {YA, TIDAK}
@attribute Bermain_Tenis {YA, TIDAK}

@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,?
```

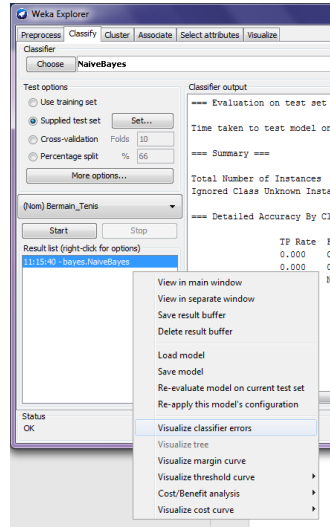
3. Buka kembali file Cuaca.arff dengan menggunakan Weka Explorer. Buka Weka Explorer – Klik Open File – Pilih Cuaca.arff



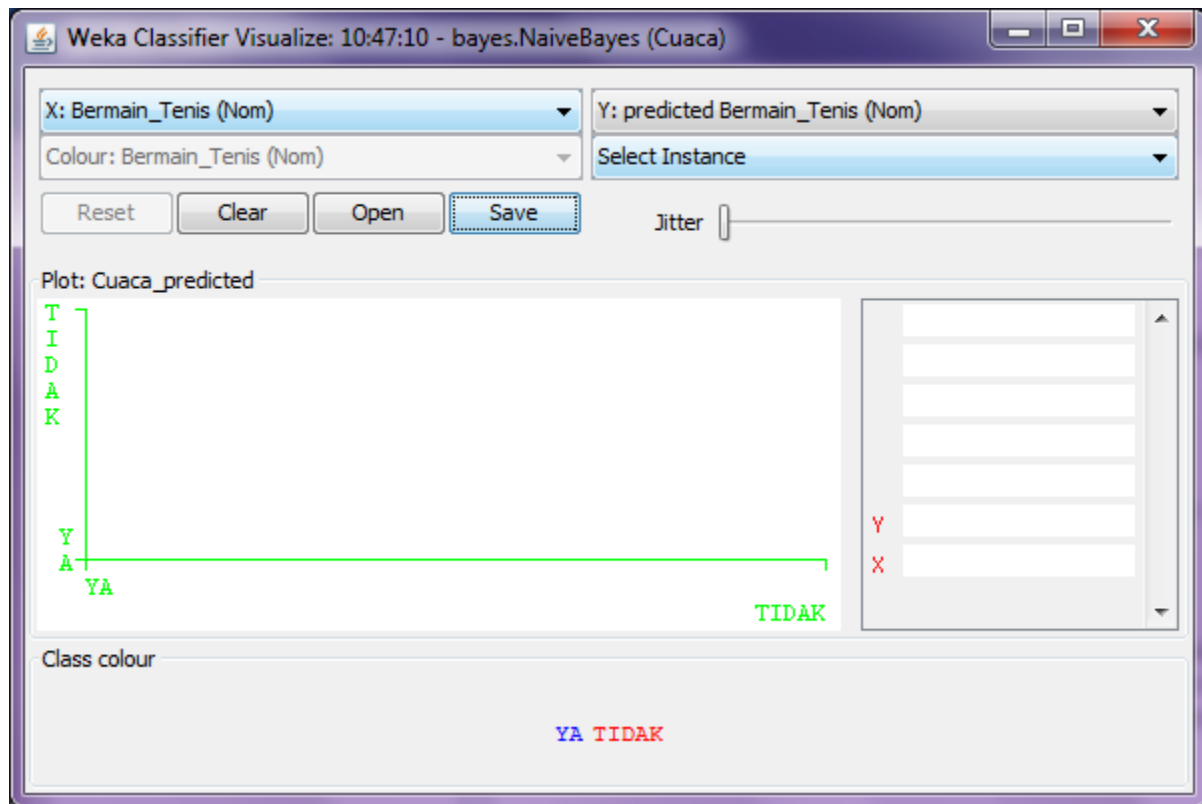
4. Masih pada Weka Explorer, pilih tab **Classify**
5. Sehingga akan muncul jendela Weka Explorer pada tab Classify, pada kotak **Classifier** klik tombol **Choose** untuk memilih metode/algoritma **Naïve Bayes**
6. Klik tombol Set untuk menentukan file ARFF sebagai data uji
7. Sehingga akan muncul jendela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

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

	A	B	C	D	E
	Cuaca	Suhu	Kelembapan Udara	Berangin	Bermain_Tenis
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah	80	90	YA	TIDAK
3	Cerah	83	86	TIDAK	YA
4	Mendung	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
14	Hujan	71	91	YA	TIDAK
15					

Training Testing (+)

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

	A	B	C	D
	Cuaca	Suhu	Kelembapan_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

Training Testing (+)

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	

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.

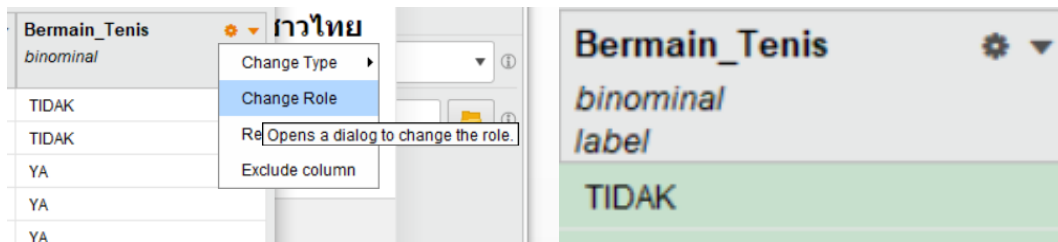
☐ Replace errors with missing values

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembaban_Udara <i>integer</i>	Berangin <i>polynomial</i>	Bermain_Tenis <i>polynomial</i>
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**



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

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

ExampleSet (//Local Repository/DataCuaca_Training)

Open in Turbo Prep Auto Model Filter (14 / 14 examples): all

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** ☒ 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**.

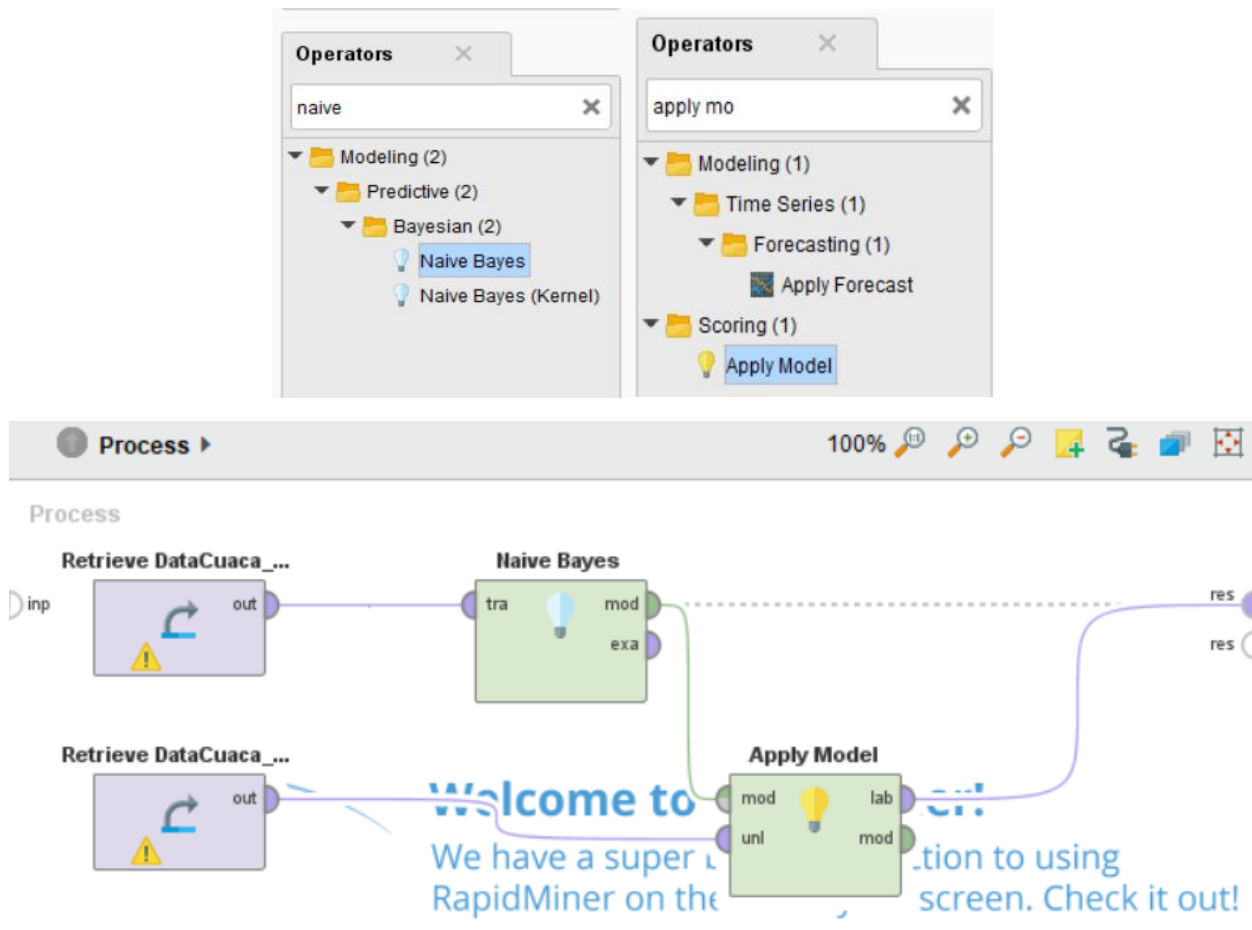
tory **ExampleSet (//Local Repository/DataCuaca_Testing)**

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 konektor 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):
✓	Prediction prediction(Bermain_Tenis)	Binominal	0	Least TIDAK (2)	Most YA (5) Values YA (5), TIDAK (2)
✓	Confidence_TIDAK confidence(TIDAK)	Real	0	Min 0.007	Max 0.856 Average 0.353
✓	Confidence_YA confidence(YA)	Real	0	Min 0.144	Max 0.993 Average 0.647
✓	Cuaca	Polynomial	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	Polynomial	0	Least TIDAK (3)	Most YA (4) Values YA (4), TIDAK (3)

Tugas

- Berdasarkan tabel berikut, buatlah file dalam format Excel (.xls) dan format ARFF (.arff)
! Data ini akan digunakan seagai **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

- 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!

a. Membuka file Training

The screenshot shows the Weka Explorer interface. The 'Preprocess' tab is active. The 'Open file...' button is highlighted. The 'Filter' section shows 'None' selected. The 'Current relation' is 'Siswa' with 20 instances and 6 attributes. The 'Attributes' list shows 'Jurusan_SMA' selected. The 'Selected attribute' section shows 'Jurusan_SMA' with 3 distinct values. The 'Class: Lama_Studi (Nom)' dropdown is set to 'Lama_Studi (Nom)'. A bar chart visualizes the distribution of 'Lama_Studi' values.

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

b. Memilih metode Naive Bayes

The screenshot shows the Weka Explorer interface with the 'Classify' tab active. The 'NaiveBayes' classifier is selected. The 'Test options' section shows 'Supplied test set' selected. The 'Classifier output' section is empty. The 'Result list' section is also empty.

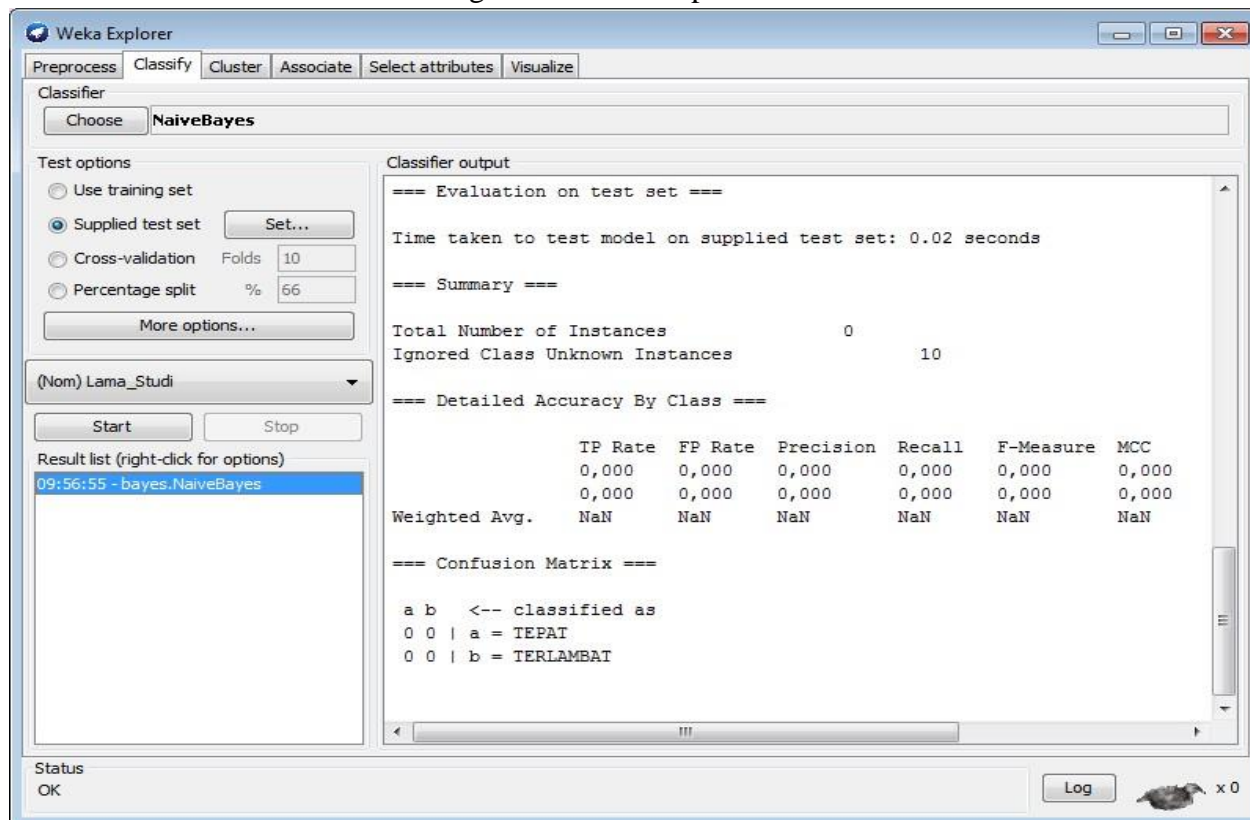
Test options:

- ☐ Use training set
- ☒ Supplied test set
- ☐ Cross-validation Folds: 10
- ☐ Percentage split %: 66

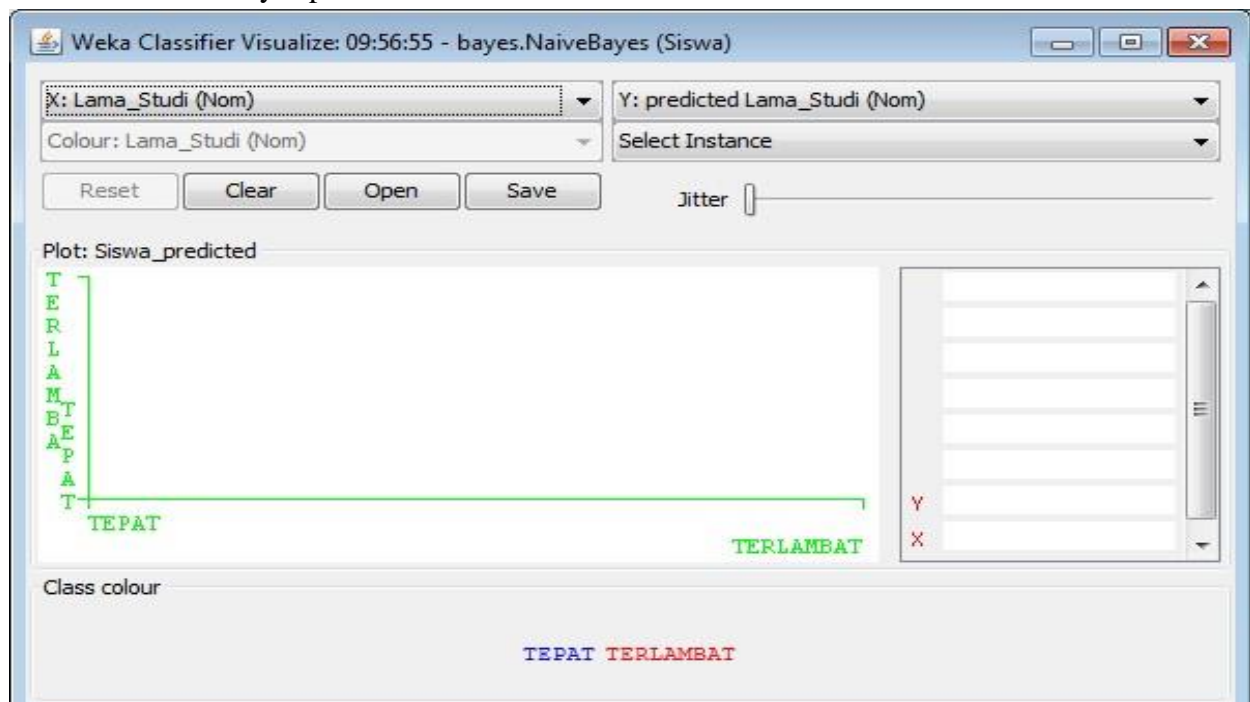
Classifier output:

Result list (right-click for options):

c. Membuka file Testing lalu klik kanan pilih Visualize classifier errors



d. Menyimpan file hasil



e. Hasil Prediksi

ARFF-Viewer - D:\dwdm\modul 8\tugas\HasiTugas.arff

File Edit View

HasiTugas.arff

Relation: Siswa_predicted

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 terhadap data testing (Excel) di atas menggunakan RapidMiner!
 - a. 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	TERLAMBAT

Previous Next Cancel

The screenshot displays the RapidMiner Studio 9.3.001 interface. The main window shows a data table with 14 rows and 7 columns. The columns are labeled: Row No., Lama_Studi, Jurusan_SMA, Gender, Asal_Sekolah, Rerata_SKS, and Asisten. The data rows show student information, with 'Lama_Studi' values ranging from 1 to 14, 'Jurusan_SMA' values including TERLAMBAT, TEPAT, and LAIN, 'Gender' values including WANITA and PRIA, 'Asal_Sekolah' values including SURAKARTA and LUAR, 'Rerata_SKS' values ranging from 16 to 20, and 'Asisten' values including TIDAK and YA. The interface includes a top menu bar (File, Edit, Process, View, Connections, Settings, Extensions, Help), a toolbar with icons for file operations and process execution, a sidebar with icons for Data, Statistics, Visualizations, and Annotations, and a right-hand pane showing the Repository with folders like Samples, DB, and Local Repository. The 'Local Repository' folder is expanded, showing sub-folders like Connections, data, and processes, and files like DataCuaca_Testing, DataCuaca_Training, DataSekolah_Training, and HasilCuaca.

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

b. 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 <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>polynomial</i>
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

← Previous
Next →
✖ Cancel

✓ no problems.

<new process> - RapidMiner Studio Educational 9.3.001 @ asus-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/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)
ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)

Result History ExampleSet (/Local Repository/DataSekolah_Testing)

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)

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9:10 AM)
 - DataCuaca_Training (asus - v1, 10/5/19 9:10 AM)
 - DataSekolah_Testing (asus - v1, 10/5/19 10:00 AM)
 - DataSekolah_Training (asus - v1, 10/5/19 10:00 AM)
 - HasilCuaca (asus - v1, 10/5/19 10:00 AM)

c. Implementasi Naive Bayes

<new process> - RapidMiner Studio Educational 9.3.001 @ asus-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 (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing (asus - v1, 10/5/19 9:10 AM)
 - DataCuaca_Training (asus - v1, 10/5/19 9:10 AM)
 - DataSekolah_Testing (asus - v1, 10/5/19 10:00 AM)
 - DataSekolah_Training (asus - v1, 10/5/19 10:00 AM)
 - HasilCuaca (asus - v1, 10/5/19 10:00 AM)

Operators

- Time Series (1)
 - Forecasting (1)
 - Apply Forecast
- Scoring (2)
 - Confidences (1)
 - Apply Threshold
 - Apply Model
- Get more operators from the Marketplace

Process

100%

Process

Retrieve DataSekola... Naive Bayes Apply Model

Parameters

Apply Model

No parameters to display.

Show advanced parameters

Change compatibility (9.3.001)

Help

Apply Model

RapidMiner Studio Core

Tags: Predict, Predictions, Forecasts, Scores, Scoring, Trained, Test

Synopsis

This Operator applies a model on an ExampleSet.

Jump to Tutorial Process

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

Activate Wisdom of Crowds

d. Hasil

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	0.868	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

Name	Type	Missing	Statistics	Filter (8 / 8 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	Polynomial	0	Least IPS (2) Most IPA (4)	Values IPA (4), LAIN (4), ...[1 more]
Gender	Polynomial	0	Least WANITA (4) Most PRIA (6)	Values PRIA (6), WANITA (4)
Asal_Sekolah	Polynomial	0	Least LUAR (3) Most SURAKARTA (7)	Values SURAKARTA (7), LUAR (3)
Rerata_SKS	Integer	0	Min 17 Max 19	Average 18.100
Asisten	Polynomial	0	Least YA (2) Most TIDAK (8)	Values TIDAK (8), YA (2)

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)	Real	0	Min 0.005	Max 0.868	Average 0.524
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)	Binominal	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. Prekdisikan 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.

a. Mengimport data Dewi

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet3 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	IPA	WANITA	LUAR 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	IPA	WANITA	LUAR SURAKARTA	18	TIDAK

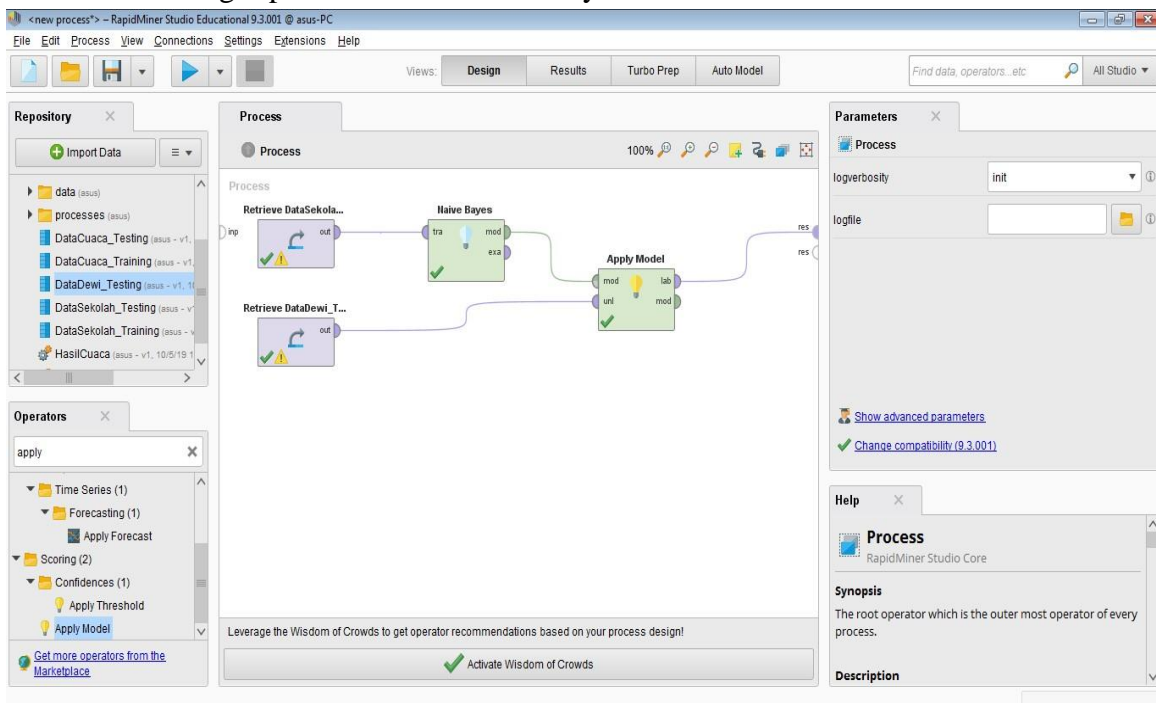
PREVI...

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

b. Mengimplementasikan Naive Bayes



c. Hasil

Open in Turbo Prep Auto Model								
Filter (1 / 1 examples): all								
Row No.	prediction(L...	confidence(...	confidence(...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.457	0.543	IPA	WANITA	LUAR SURA...	18	TIDAK

RapidMiner Studio Educational 9.3.001 @ asus-PC

Views: Design Results Turbo Prep Auto Model

Filter (8 / 8 attributes): Search for Attributes

Name	Type	Missing	Statistics		
prediction(Lama_studi)	Binominal	0	Least TERLAMBAT (0)	Most TEPAT (1)	Values TEPAT (1), TERLAMBAT (0)
confidence(TERLAMBAT)	Real	0	Min 0.457	Max 0.457	Average 0.457
confidence(TEPAT)	Real	0	Min 0.543	Max 0.543	Average 0.543
Jurusan_SMA	Polynominal	0	Least IPA (1)	Most IPA (1)	Values IPA (1)
Gender	Polynominal	0	Least WANITA (1)	Most WANITA (1)	Values WANITA (1)
Asal_Sekolah	Polynominal	0	Least LUAR SURAKARTA (1)	Most LUAR SURAKARTA (1)	Values LUAR SURAKARTA (1)
Rerata_SKS	Integer	0	Min 18	Max 18	Average 18
Asisten	Polynominal	0	Least TIDAK (1)	Most TIDAK (1)	Values TIDAK (1)

Showing attributes 1 - 8 Examples: 1 Special Attributes: 3 Regular Attributes: 5

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.

a. Mengimport Data Jono

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Sheet4 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	PRIA	SURAKARTA	17.000	YA

Previous Next Cancel

ExampleSet (1 example, 0 special attributes, 5 regular attributes)

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

b. Mengimplementasikan Naive Bayes

Process flow diagram showing the implementation of Naive Bayes:

- Retrieve DataSekola...** (Input)
- Retrieve DataJono_T...** (Input)
- Naive Bayes** (Model)
- Apply Model** (Output)

Parameters for the Naive Bayes process:

- logverbosity: init
- logfile: [empty]

Help for the Naive Bayes process:

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

Description

c. Hasil

ExampleSet (1 example, 0 special attributes, 5 regular attributes)

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

ExampleSet (/Local Repository/DataCuaca_Testing) ExampleSet (/Local Repository/DataCuaca_Training)
ExampleSet (/Local Repository/DataSekolah_Testing) ExampleSet (/Local Repository/DataSekolah_Training)
ExampleSet (/Local Repository/DataDewi_Testing) ExampleSet (/Local Repository/DataSekolah_Testing)
ExampleSet (/Local Repository/DataJono_Testing)

Result History ExampleSet (Apply Model)

Name	Type	Missing	Statistics			Filter (8 / 8 attributes)
Prediction prediction(Lama_Studi)	Binominal	0	Least TERLAMBAT (0)	Most TEPAT (1)	Values TEPAT (1), TERLAMBAT (0)	
Confidence_TERLAMBAT confidence(TERLAMBAT)	Real	0	Min 0.076	Max 0.076	Average 0.076	
Confidence_TEPAT 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)	
Rerata_SKS	Integer	0	Min 17	Max 17	Average 17	
Asisten	Polynomial	0	Least YA (1)	Most YA (1)	Values YA (1)	

Showing attributes 1 - 8 Examples: 1 Special Attributes: 3 Regular Attributes: 5

Repository

- Import Data
- Samples
- DB (Legacy)
- Local Repository (asus)
 - Connections (asus)
 - data (asus)
 - processes (asus)
 - DataCuaca_Testing
 - DataCuaca_Training
 - DataDewi_Testing
 - DataJono_Testing
 - DataSekolah_Testing
 - DataSekolah_Training
 - HasilCuaca (asus - v1)
 - HasilDewi (asus - v1)
 - HasilSekolah (asus - v1)

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