

Nama : Muhammad Khotibul Umam Senoaji

NIM : L200170050

Kelas : C

+-----+

MODUL VII

KLASIFIKASI : NAÏVE BAYES

8.4 Langkah-langkah Praktikum

8.4.1 Implementasi Naïve Bayes dengan Weka

1. Membuat table Data Testing Cuaca yang ada dimodul ke bentuk ARFF dan simpan dengan nama **CuacaTesting.arff**.

```
1  @relation Cuaca
2
3  @attribute Cuaca {Cerah, Mendung, Hujan}
4  @attribute Suhu real
5  @attribute Kelembaban_udara real
6  @attribute Barangin {YA, TIDAK}
7  @attribute Bermain_Tenis {YA, TIDAK}
8
9  @data
10 Cerah,75,65,TIDAK,?
11 Cerah,80,68,YA,?
12 Cerah,83,87,YA,?
13 Mendung,70,96,TIDAK,?
14 Mendung,68,81,TIDAK,?
15 Hujan,65,75,YA,?
16 Hujan,64,85,YA,?
```

2. Membuka file **Cuaca.arff** pada tab **Preprocess**, lalu pada tab **Classify** pilih mode **Naïve Bayes**, klik **Supplied test set**, klik **set**, klik **Open file...** lalu pilih file **CuacaTesting.arff**, **Close**, lalu klik **Start**.

The screenshot shows the Weka Explorer application window. The 'Preprocess' tab is active, displaying the 'Cuaca' dataset with 5 attributes and 14 instances. The 'Classify' tab is also visible. The 'Selected attribute' section shows 'Cuaca' with 3 distinct values and a nominal type. A table lists the values: 'Cerah' (5), 'Mendung' (4), and 'Hujan' (5). The 'Class' dropdown is set to 'Bermain_Tenis (Nom)'. A bar chart visualization shows the distribution of the 'Bermain_Tenis' class across the 'Cuaca' categories, with red bars for 'Cerah' and 'Hujan' and a blue bar for 'Mendung'.

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter: Choose **None** Apply

Current relation
Relation: Cuaca
Instances: 14
Attributes: 5
Sum of weights: 14

Attributes
All | None | Invert | Pattern

No.	Name
1	<input checked="" type="checkbox"/> Cuaca
2	<input type="checkbox"/> Suhu
3	<input type="checkbox"/> Kelembaban_udara
4	<input type="checkbox"/> Barangin
5	<input type="checkbox"/> Bermain_Tenis

Remove

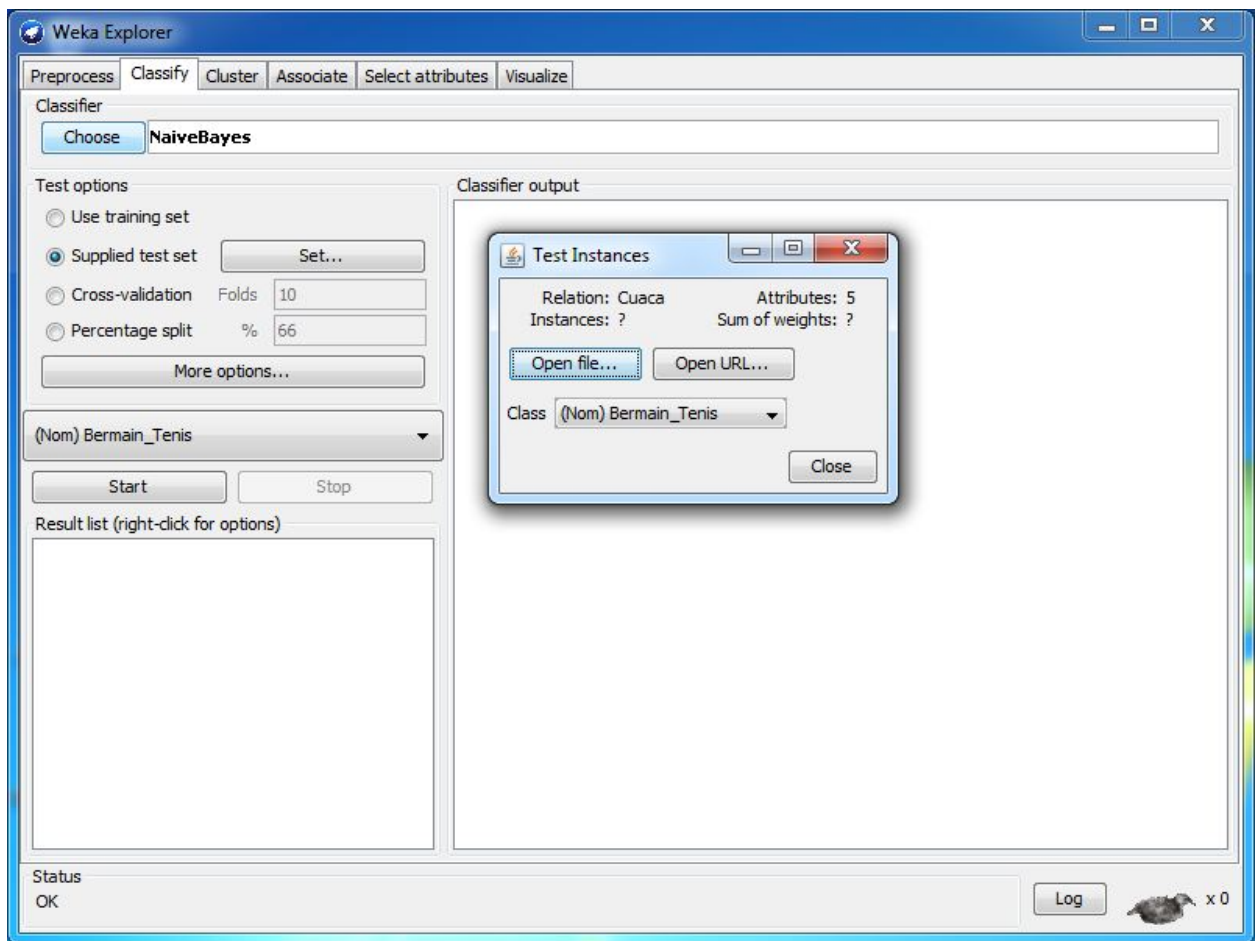
Selected attribute
Name: Cuaca
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

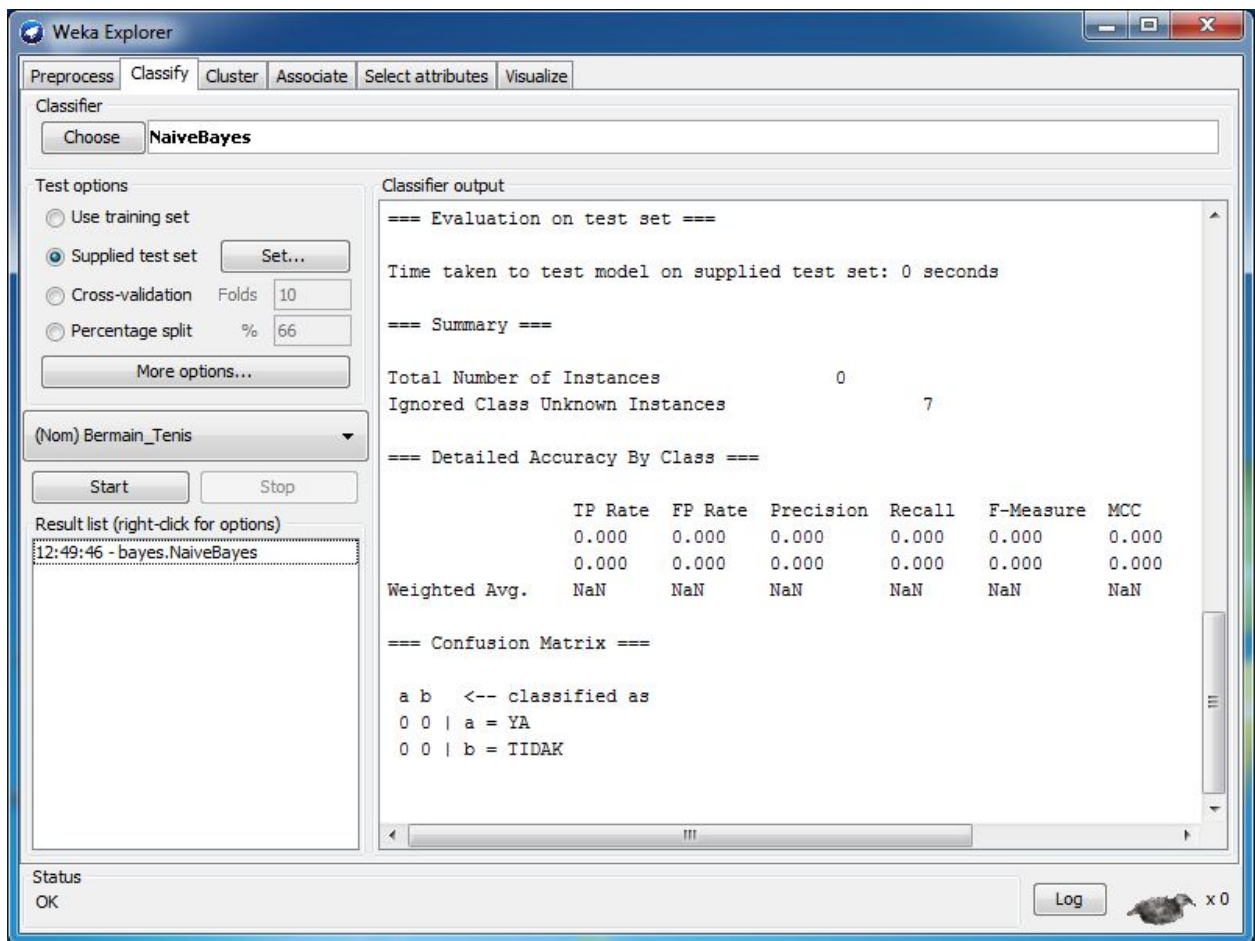
No.	Label	Count	Weight
1	Cerah	5	5.0
2	Mendung	4	4.0
3	Hujan	5	5.0

Class: Bermain_Tenis (Nom) Visualize All

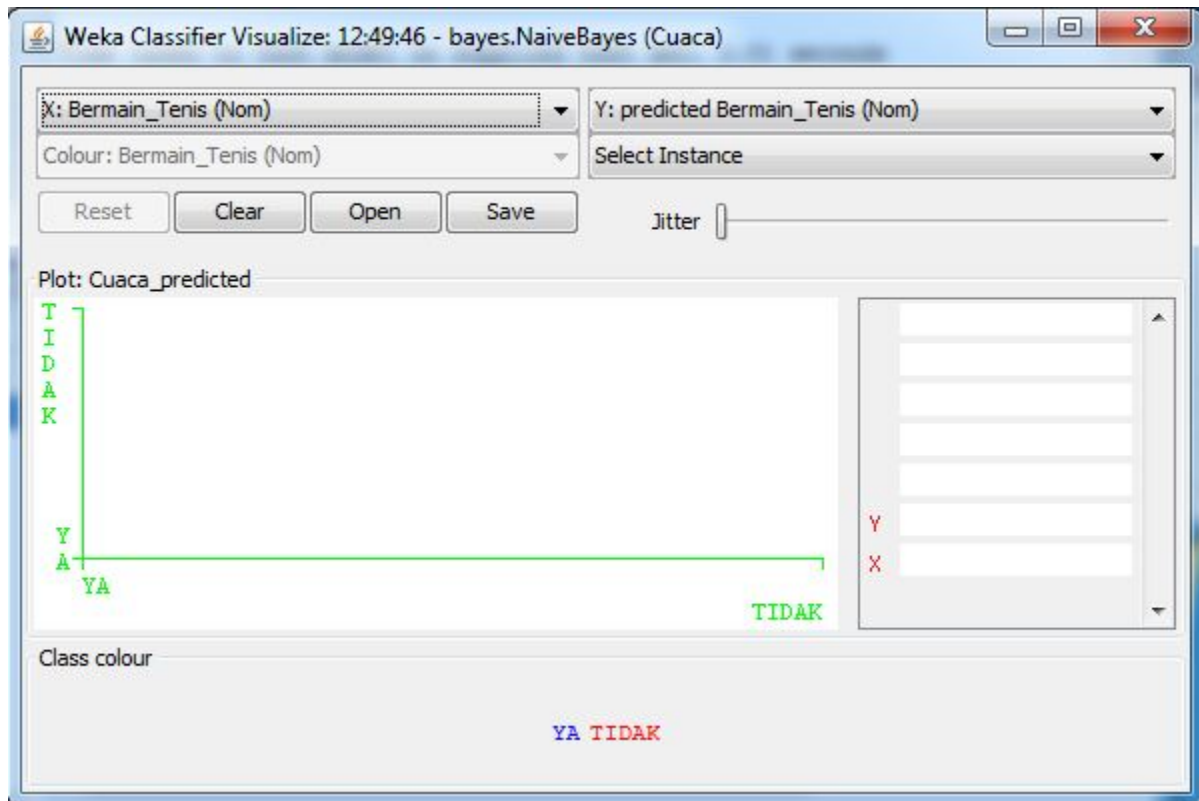
5 4 5

Status
OK Log x 0





3. Klik kanan pada **Result list**, dan pilih opsi **Classifier Visualize**, biarkan apa yang ditampilkan dan klik **save**.



4. Kembali ke GUI utama **WEKA** dan pilih opsi **Tool** => **ArffViewer** => **File** => **Open** => pilih file hasil dari **Classifier Visualize (HasilPrediksi.arff)** untuk menampilkan hasil prediksi yang sudah dibuat.

ARFF-Viewer - F:\Tugas\Praktikum Data Warehousing dan Data Mining\Pertemuan Ke-05\HasilPrediksi.arff

File Edit View

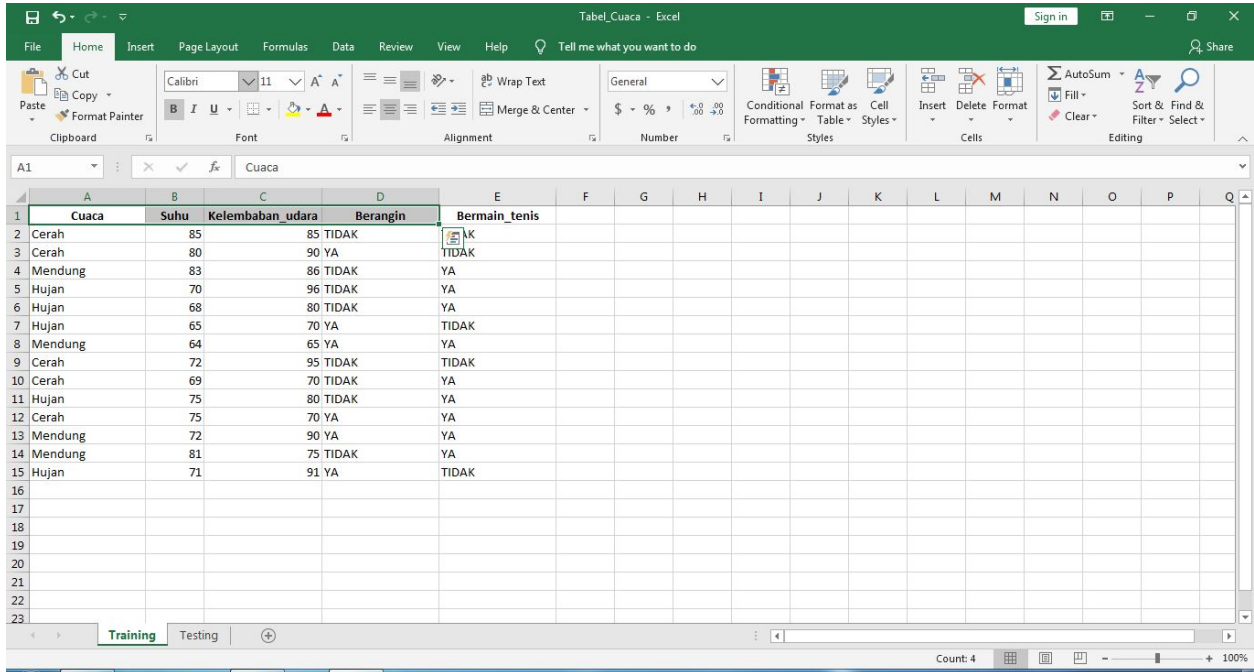
HasilPrediksi.arff

Relation: Cuaca_predicted

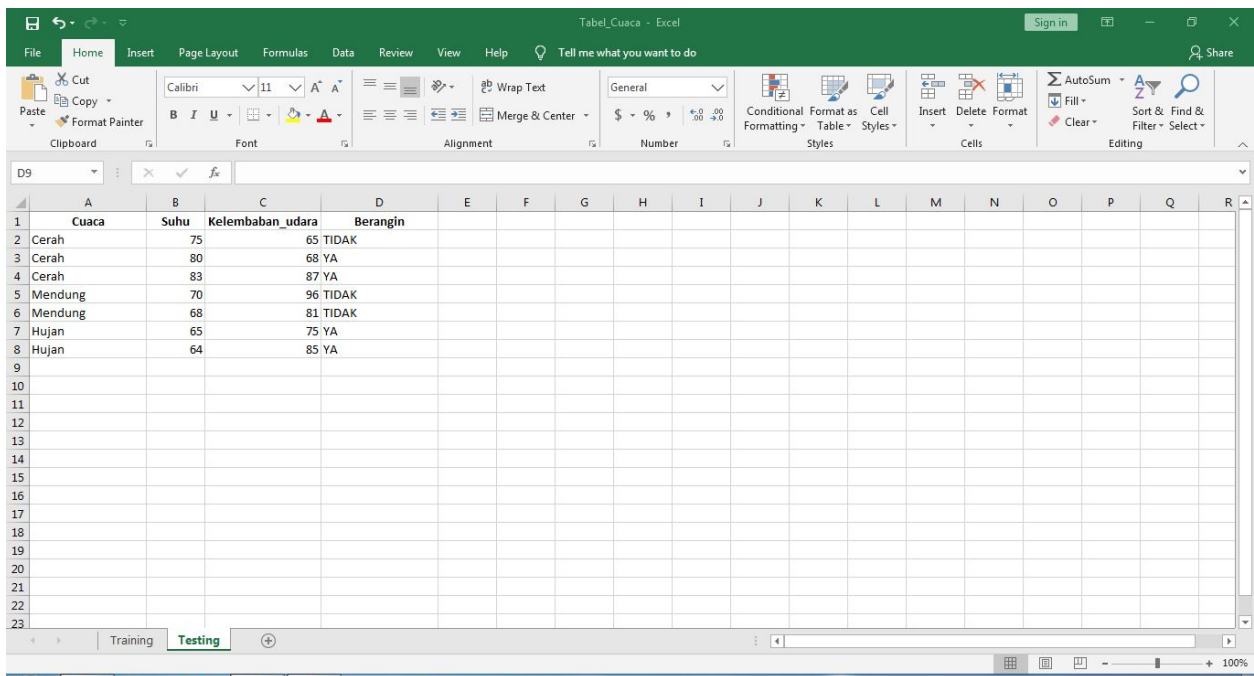
No.	1: Cuaca Nominal	2: Suhu Numeric	3: Kelembaban_udara Numeric	4: Barangin Nominal	5: prediction margin Numeric	6: predicted Bermain_Tenis Nominal	7: Bermain_Tenis Nominal
1	Cerah	75.0	65.0	TIDAK	0.762765	YA	
2	Cerah	80.0	68.0	YA	0.087878	YA	
3	Cerah	83.0	87.0	YA	-0.676866	TIDAK	
4	Mendung	70.0	96.0	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	

8.4.2 Implementasi Naïve Bayes dengan RapidMiner

1. Mempersiapkan data dengan **Microsoft Excel** dengan membuat table **Training** dan **Testing** dengan tema **Cuaca**.

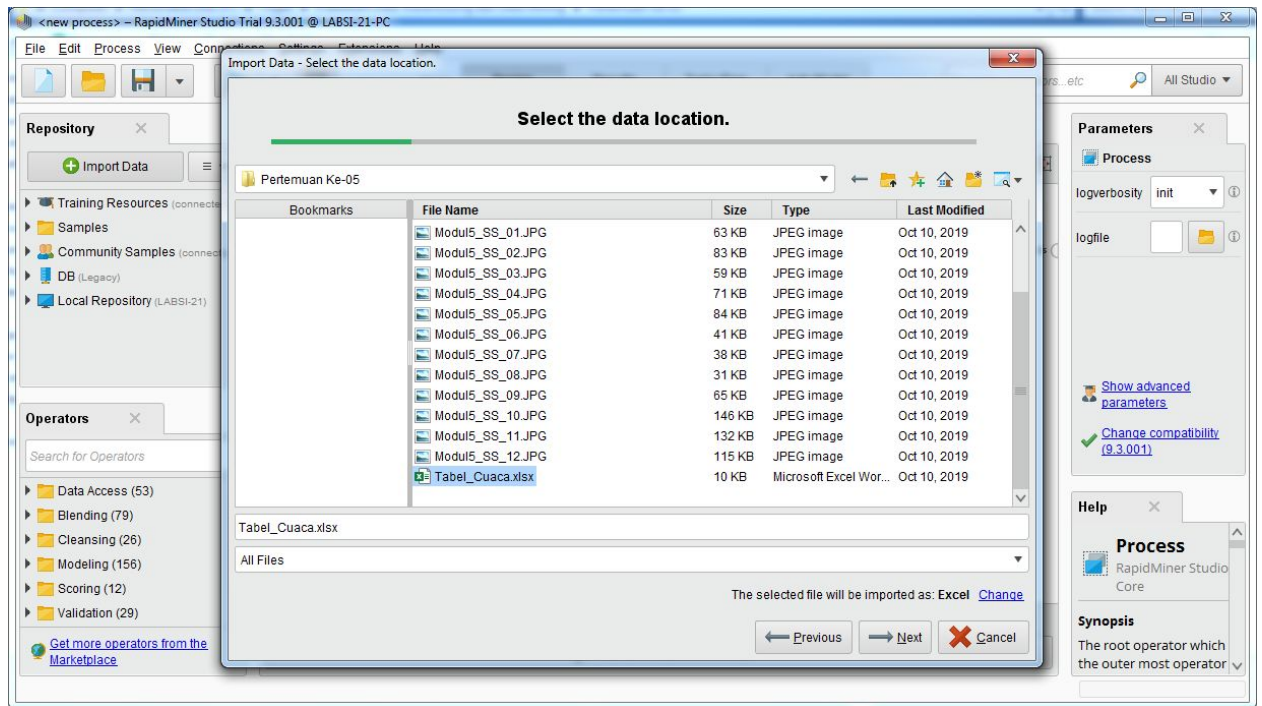


	A	B	C	D	E
1	Cuaca	Suhu	Kelembaban_udara	Berangin	Bermain_tenis
2	Cerah	85	85	TIDAK	TIDAK
3	Cerah	80	90	YA	TIDAK
4	Mendung	83	86	TIDAK	YA
5	Hujan	70	96	TIDAK	YA
6	Hujan	68	80	TIDAK	YA
7	Hujan	65	70	YA	TIDAK
8	Mendung	64	65	YA	YA
9	Cerah	72	95	TIDAK	TIDAK
10	Cerah	69	70	TIDAK	YA
11	Hujan	75	80	TIDAK	YA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



	A	B	C	D
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

2. Menjalankan **Rapid Miner** dan Meng-import file excel yang sudah disimpan tadi (nama file **Tabel_Cuaca.xls**).



3. Setelah itu pilih **Sheet** ke **Training**, klik **Next** lalu ubah tipe table **Bermain_Tenis** data dari *polynomial* dengan *binomial* dan mengganti **Role** dengan label, setelah itu simpan data ke **Local Repository**.

Import Data - Select the cells to import.

Select the cells to import.

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

	A	B	C	D	E
1	Cuaca	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	74.000	84.000	YA	TIDAK

Previous Next Cancel

Format your columns.

☐ Replace errors with missing values ⓘ

	Cuaca <i>polynomial</i>	Suhu <i>integer</i>	Kelembaban_u... <i>integer</i>	Berangin <i>polynomial</i>	Bermain_tenis <i>binominal label</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
8	Cerah	72	95	TIDAK	TIDAK
9	Cerah	69	70	TIDAK	YA
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA
13	Mendung	81	75	TIDAK	YA

✓ no problems.

← Previous

→ Next

✗ Cancel

4. Lakukan langkah yang sama pada **Sheet Testing**, dengan perbedaan tidak ada tabel yang diubah **Role** menjadi **label**.

Import Data - Select the cells to import.

Select the cells to import.

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

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

← Previous → Next X Cancel

Format your columns.☐ Replace errors with missing values ⓘ

	Cuaca <i>polynominal</i> ⚙️ ▼	Suhu <i>integer</i> ⚙️ ▼	Kelembaban_udara <i>integer</i> ⚙️ ▼	Berangin <i>binominal</i> ⚙️ ▼
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

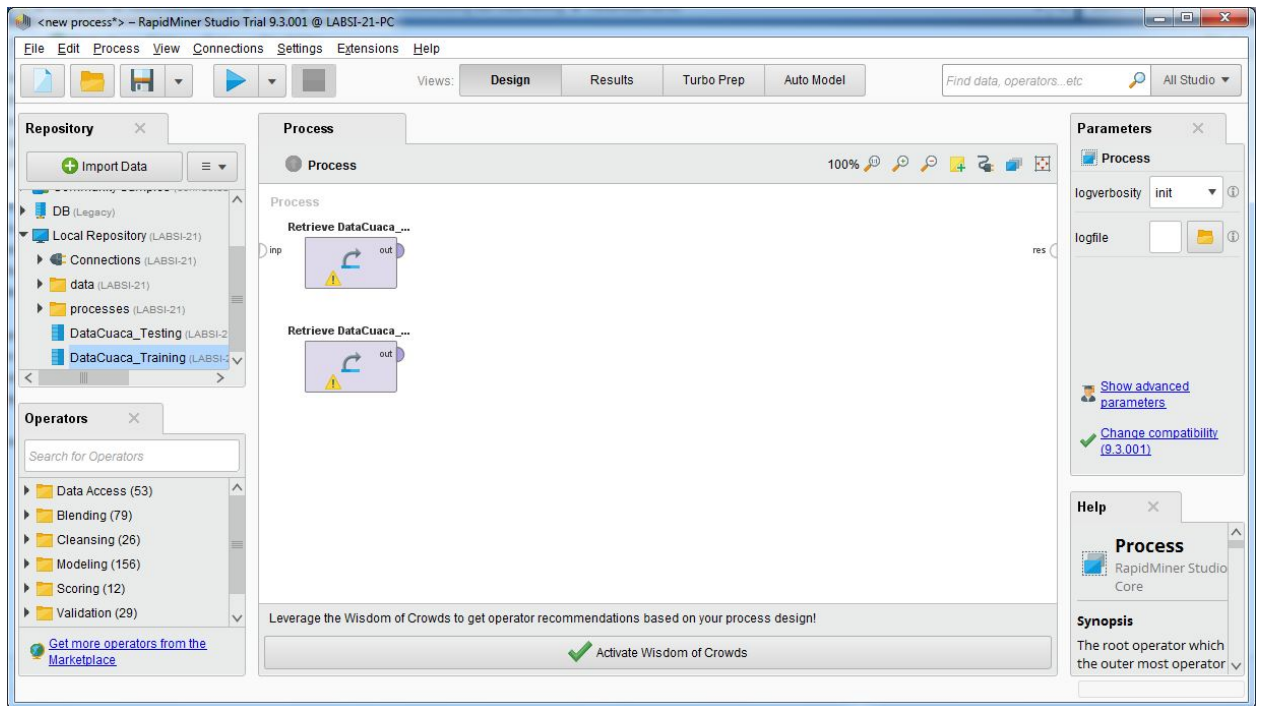
✔️ no problems.

⬅️ Previous

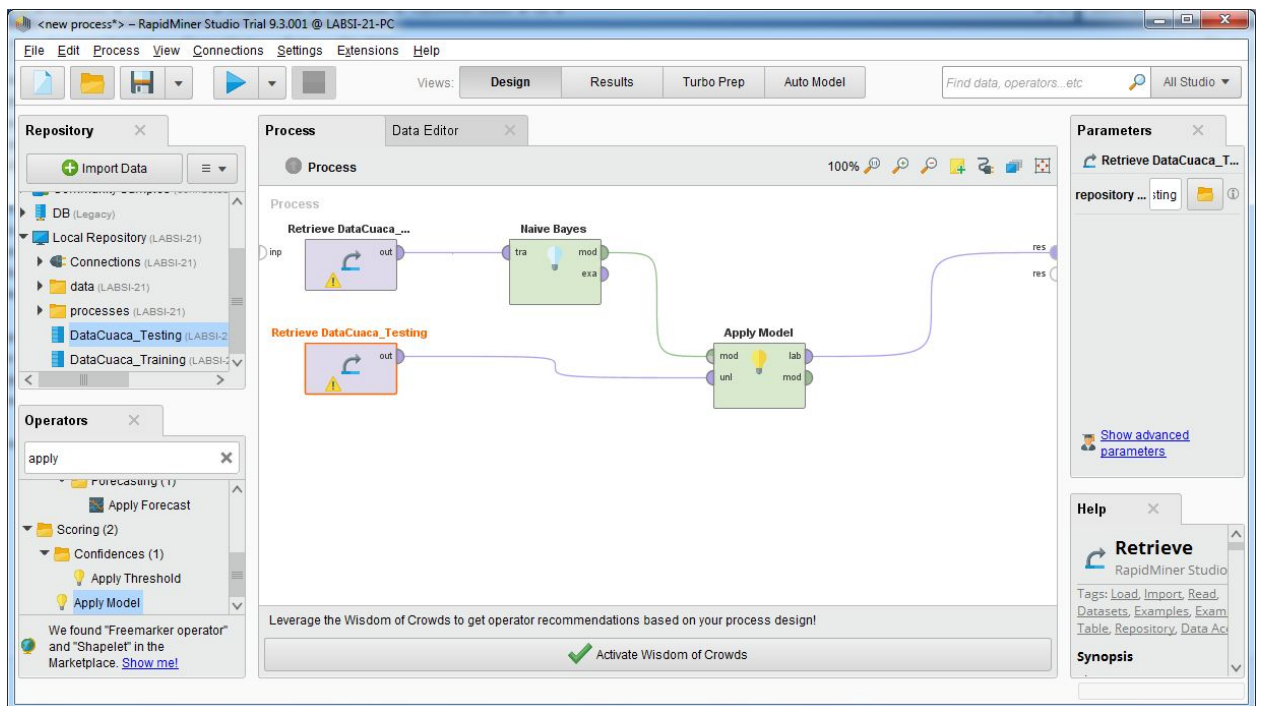
➡️ Next

❌ Cancel

5. Kembali ke halaman utama **Rapid Miner** dan membuat desain **Naïve Bayes**, dengan mengdrag **DataCuaca_Training** dan **DataCuaca_Testing** ke dalam jendela **Process View**.



6. Cari operator **Naïve Bayes** dan **Apply Model** (bisa dicari di jendela **Operator**) dan drag ke jendela **Process**, lalu sambungkan masing-masing node, lalu klik simbol **Run**.



<new process*> - RapidMiner Studio Trial 9.3.001 @ LABSI-21-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/DataCuaca_Training)

Result History

ExampleSet (Apply Model)

ExampleSet (/Local Repository/DataCuaca_Testing)

Open in Turbo Prep Auto Model

Filter (7 / 7 examples): all

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)

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

DB (Legacy)

Local Repository (LABSI-21)

Connections (LABSI-21)

data (LABSI-21)

processes (LABSI-21)

DataCuaca_Testing (LABSI-21 - v1)

DataCuaca_Training (LABSI-21 - v1)

7. Klik **Statistics** untuk melihat data lebih detail.

<new process*> - RapidMiner Studio Trial 9.3.001 @ LABSI-21-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_Training)

ExampleSet (/Local Repository/DataCuaca_Testing)

ExampleSet (/Local Repository/DataCuaca_Training)

ExampleSet (/Local Repository/DataCuaca_Testing)

Result History

ExampleSet (Apply Model)

Name Type Missing Statistics Filter (7 / 7 attributes): Search for Attributes

Name	Type	Missing	Statistics	Filter (7 / 7 attributes):	Search for Attributes
✓ prediction(Bermain_tenis)	Binominal	0	Least TIDAK (2)	Most YA (5)	Values YA (5).
✓ 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).
✓ Suhu	Integer	0	Min 64	Max 83	Average 72.143
✓ Kelembaban_udara	Integer	0	Min 65	Max 96	Average 79.571
✓ Berangin	Binominal	0	Least TIDAK (3)	Most YA (4)	Values YA (4).

Showing attributes 1 - 7

Examples: 7 Special Attributes: 3 Regular Attributes: 4

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

DB (Legacy)

Local Repository (LABSI-21)

Connections (LABSI-21)

data (LABSI-21)

processes (LABSI-21)

DataCuaca_Testing (LABSI-21 - v1, 10/10/19 1:42)

DataCuaca_Training (LABSI-21 - v1, 10/10/19 1:42)