

LAPORAN PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

MODUL 8 “KLASIFIKASI: NAIVE BAYES”



Oleh:

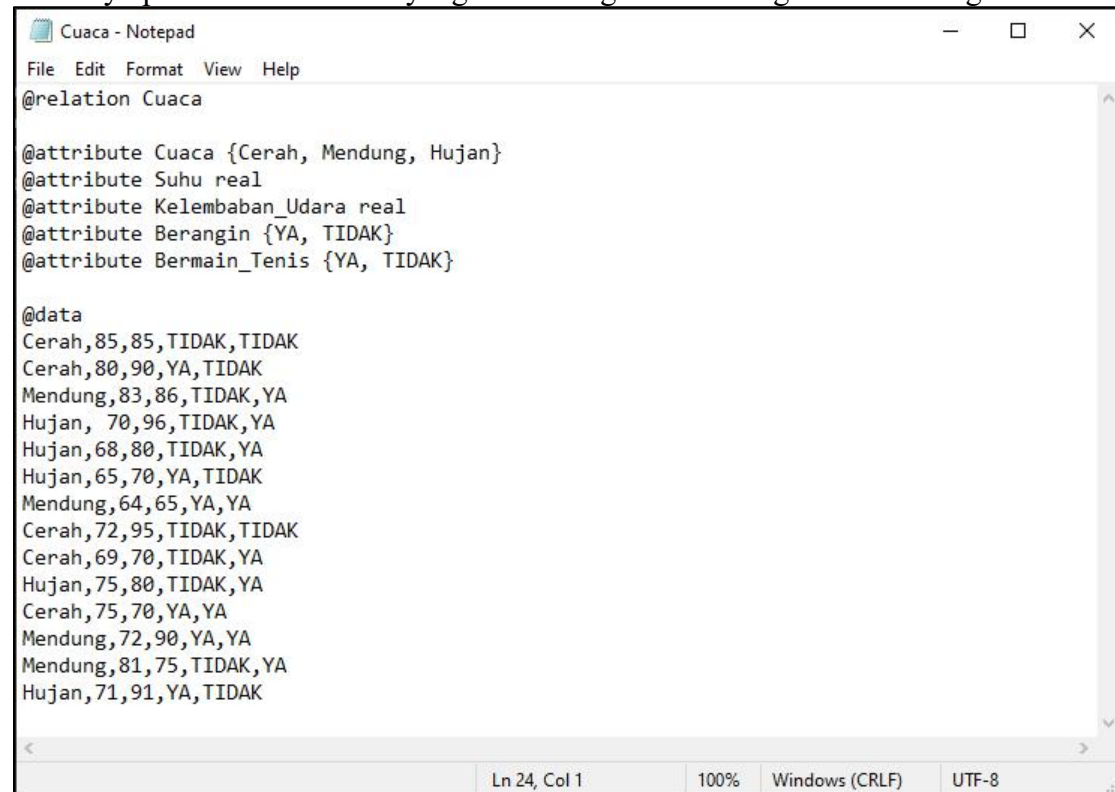
NAMA	: Daffa Putra Alwansyah
NIM	: L200190031
KELAS	: B
PRODI	: INFORMATIKA

**Fakultas Komunikasi dan Informatika
Universitas Muhammadiyah Surakarta**

Langkah-Langkah Praktikum

➤ Implementasi Naïve Bayes dengan Weka

1. Menyiapkan file Cuaca.arff yang akan kita gunakan sebagai data training.



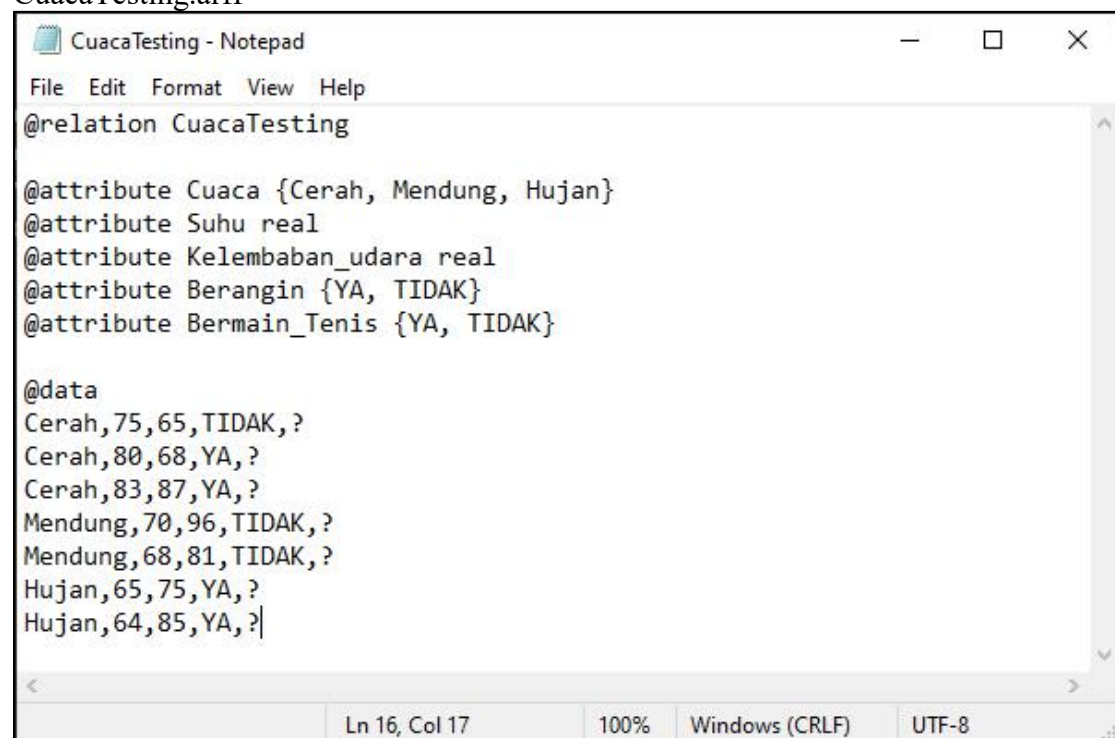
```
Cuaca - Notepad
File Edit Format View Help
@relation Cuaca

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

@data
Cerah,85,85,TIDAK,TIDAK
Cerah,80,90,YA,TIDAK
Mendung,83,86,TIDAK,YA
Hujan, 70,96,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,75,80,TIDAK,YA
Cerah,75,70,YA,YA
Mendung,72,90,YA,YA
Mendung,81,75,TIDAK,YA
Hujan,71,91,YA,TIDAK

Ln 24, Col 1    100%    Windows (CRLF)    UTF-8
```

2. Membuat data testing sebagai data uji yang akan diprediksi dengan memiliki variable-variable independent dan variable dependen yang sama. Diberi nama CuacaTesting.arff



```
CuacaTesting - Notepad
File Edit Format View Help
@relation CuacaTesting

@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,?

Ln 16, Col 17    100%    Windows (CRLF)    UTF-8
```

3. Buka kembali file Cuaca.arff dengan aplikasi weka

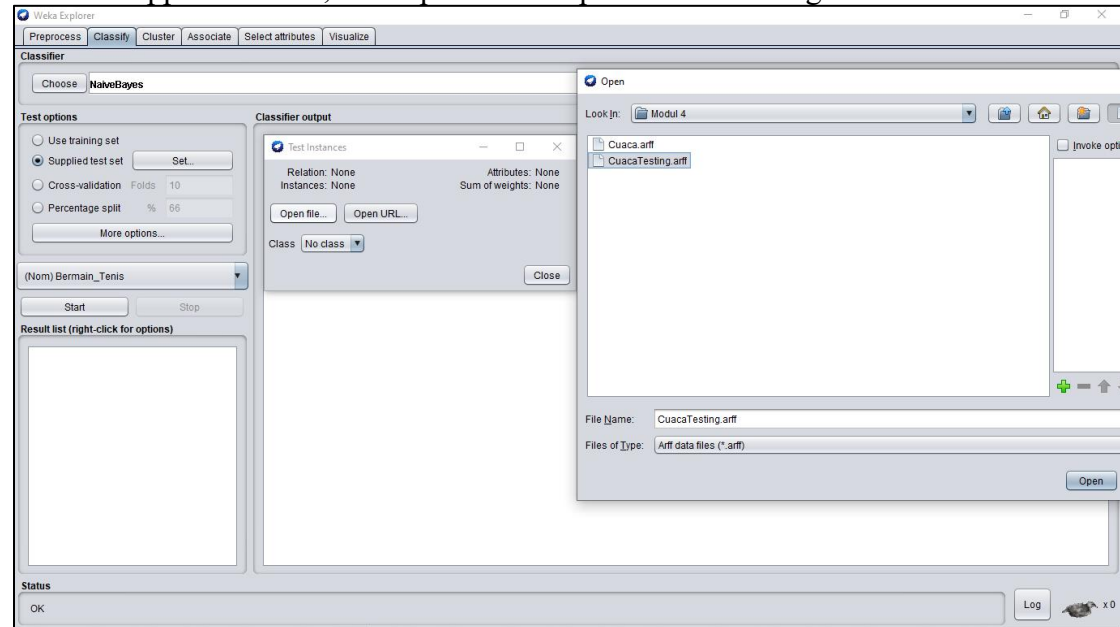
The screenshot shows the Weka Explorer interface. The 'Current relation' panel displays the 'Cuaca' dataset with 14 instances and 5 attributes. The 'Selected attribute' panel shows the 'Bermain_Tenis' attribute with 3 distinct values and a nominal type. The 'Attributes' panel lists the attributes: Cuaca, Suhu, Kelembaban_Udara, Berangin, and Bermain_Tenis. The 'Visualize All' button is visible at the bottom right.

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

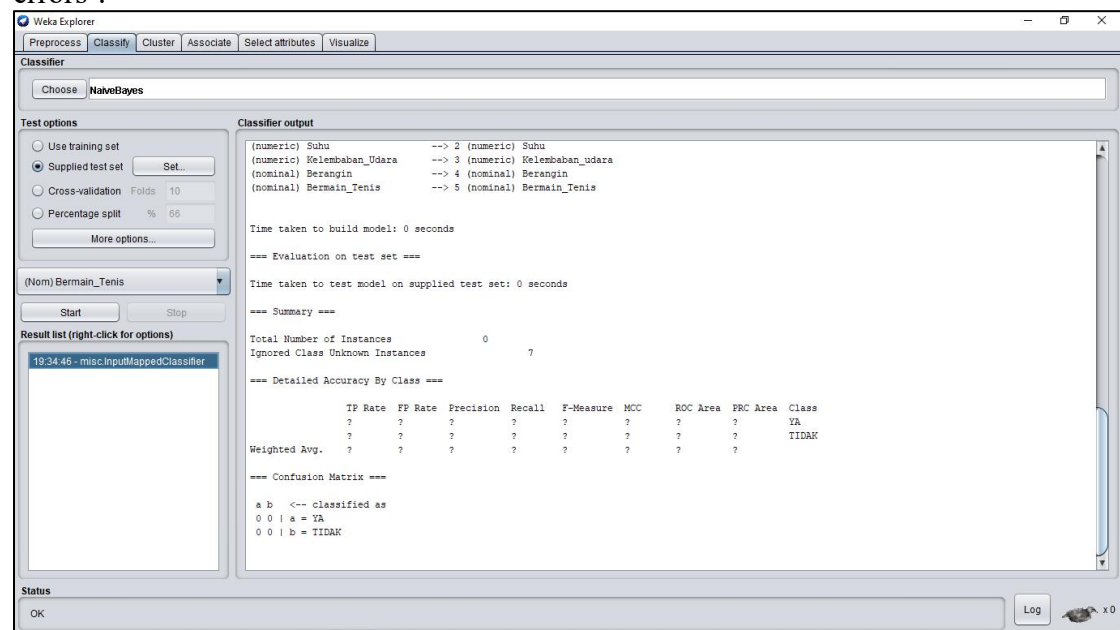
4. Setelah itu pilih tab Classify, kemudian klik Choose dan pilih NaiveBayes.

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The 'Choose' dialog box is open, displaying a tree view of classifiers. The 'NaiveBayes' classifier is selected under the 'bayes' category. The 'Close' button is visible at the bottom right of the dialog box.

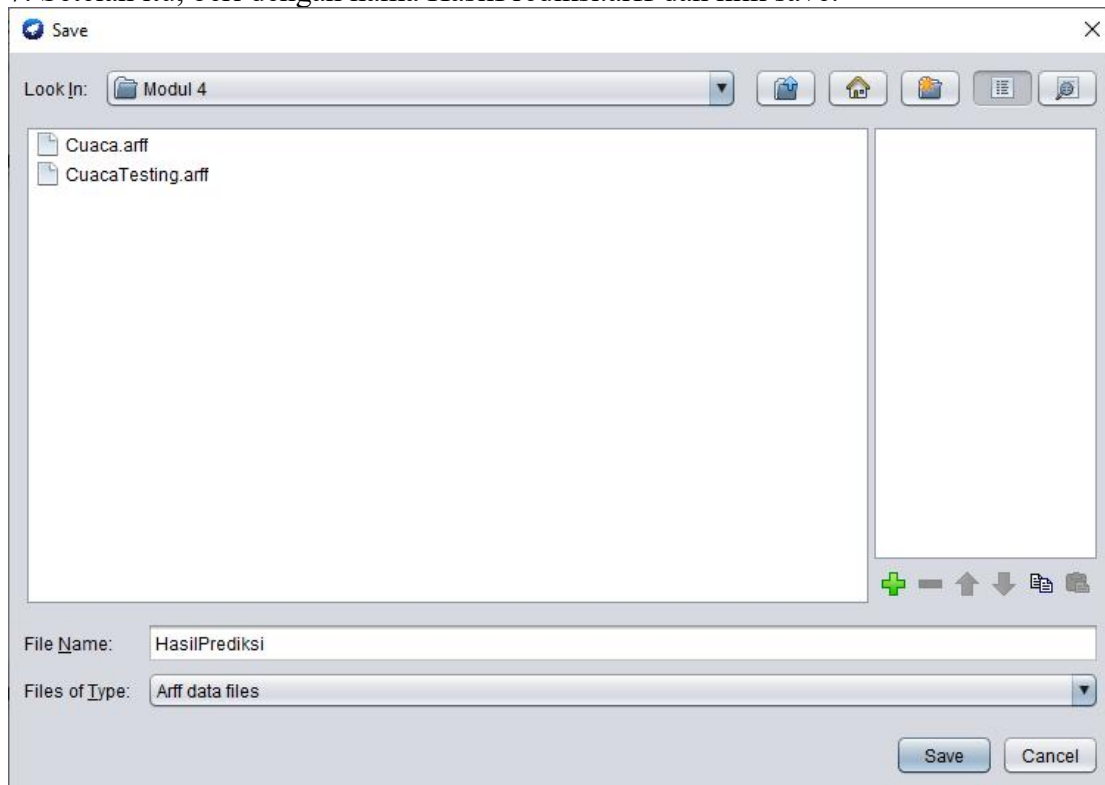
5. Pilih Supplied test set, klik Open file dan pilih CuacaTesting.arff.



6. Class yang dipilih adalah Bermain_Tenis, setelah itu klik Start untuk memulai proses naïve bayes, setelah itu klik kanan pada result, dan pilih 'visualize classifier errors'.



7. Setelah itu, beri dengan nama HasilPrediksi.arff dan klik save.



8. Tutup Weka Exploer, kembali ke menu utama, pilih Tool > ArffViewer, buka file HasilPrediksi.arff yang sudah disimpan tadi.

No.	Cuaca	Suhu	Kelembaban_udara	Berangin	prediction margin	predicted Bermain_Tenis	Bermain_Tenis
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	Mend...	70.0	96.0	TIDAK	0.628523	YA	
5	Mend...	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. Mempersiapkan file Tabel_Cuaca.xls yang terdiri dari 2 sheet, sheet pertama berupa Training dan sheet kedua berupa Testing. Sheet 1 digunakan sebagai data training dan sheet2 digunakan sebagai data uji.

WPS Office Tabel_Cuaca.xls

Menu Paste Cut Copy Format Painter Calibri 11 Home Insert Page Layout Formulas Data Review View Tools Click to find commands

General Conditional Formatting Format as Table AutoSum AutoFilter Sort Fill Format

Merge and Center Wrap Text

fx

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
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																			

Training Testing +

WPS Office Tabel_Cuaca.xls

Menu Paste Copy Format Painter Calibri 11 Bold Italic Underline Text Color Background Color Merge and Center Wrap Text General Conditional Formatting Format as Table AutoSum AutoFilter Sort Fill Format

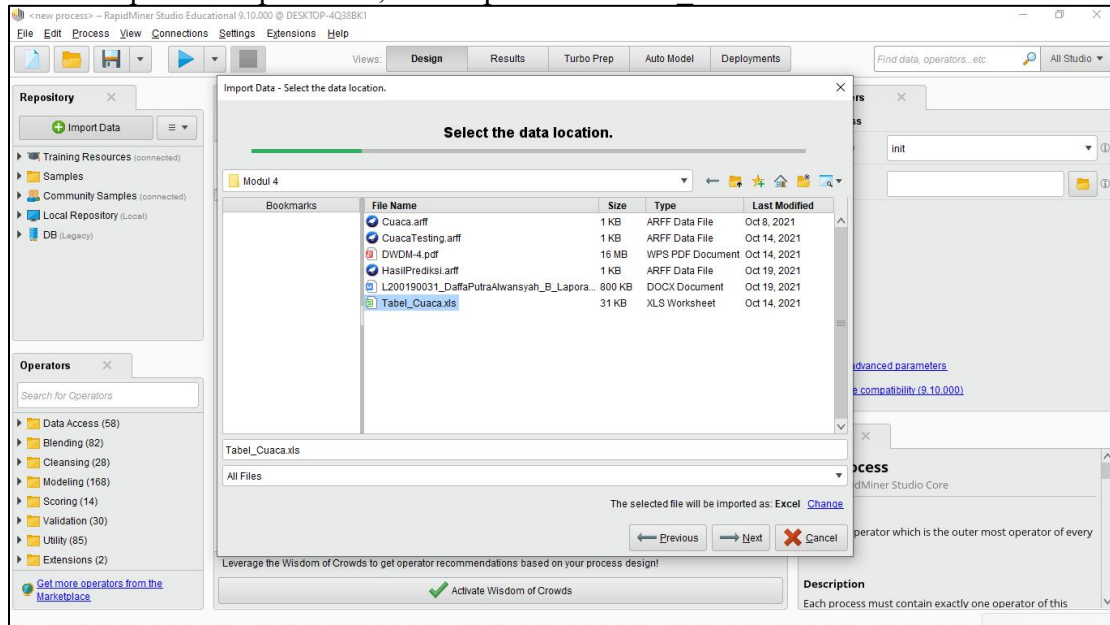
Click to find commands

A9 fx

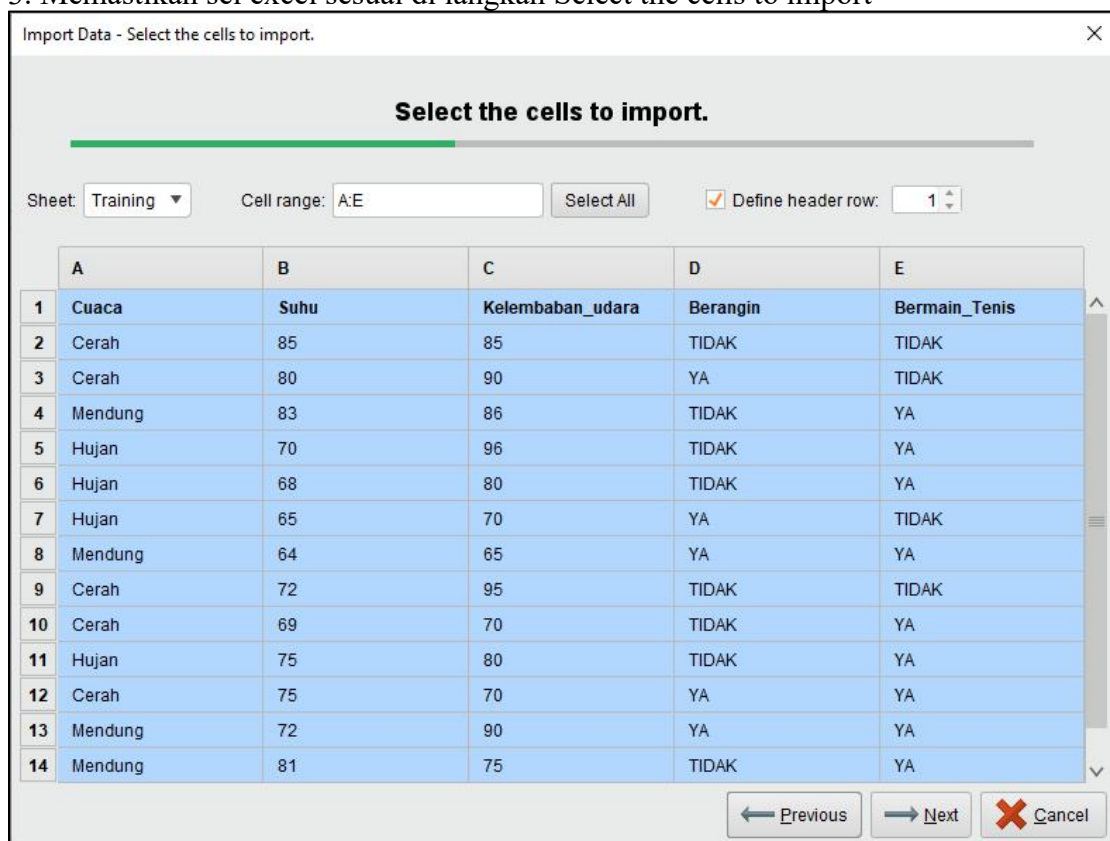
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Cuaca	Suhu	Kelembaban	udara	Berangin															
2	Cerah	75		65	TIDAK															
3	Cerah	80		68	YA															
4	Cerah	83		87	YA															
5	Mendung	70		96	TIDAK															
6	Mendung	68		81	TIDAK															
7	Hujan	65		75	YA															
8	Hujan	64		85	YA															
9																				
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				
19																				
20																				
21																				
22																				
23																				
24																				
25																				
26																				

Training Texting

2. Buka aplikasi RapidMiner, lalu import data Tabel Cuaca.xls



3. Memastikan sel excel sesuai di langkah Select the cells to import



4. Setelah itu ubah type Berangin dan Bermain_Tenis menjadi 'binominal' dan ubah role Bermain_Tenis menjadi label.

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Cuaca <i>polynominal</i>	Suhu <i>integer</i>	Kelembaban_u... <i>integer</i>	Berangin <i>binominal</i>	Bermain_Tenis <i>binominal</i>
1	Cerah	85	85	TIDAK	TIDAK
2	Cerah				AK
3	Mendung				
4	Hujan				
5	Hujan				
6	Hujan				AK
7	Mendung				AK
8	Cerah				
9	Cerah				
10	Hujan	75	80	TIDAK	YA
11	Cerah	75	70	YA	YA
12	Mendung	72	90	YA	YA

Change role

Please enter the new role:

label

OK Cancel

no problems.

Previous Next Cancel

5. Simpan dengan nama DataCuaca_Training.

Import Data - Where to store the data?

Where to store the data?

Local Repository (Local)

- data
- processes

Name DataCuaca_Training

Location //Local Repository/DataCuaca_Training

Previous Finish Cancel

6. Menampilkan hasil import file Tabel_Cuaca.xls pada sheet1t file Tabel_Cuaca.xls pada sheet1(Training).

The screenshot shows the RapidMiner Studio interface. The 'Results' tab is active, displaying a table with 14 examples of weather data. The table has columns: Row No., Bermain_Te..., Cuaca, Suhu, Kelembaban..., and Berangin. The data is as follows:

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)

7. Ulangi hal yang sama pada sheet1 tadi, tapi disini Sheet menggunakan Testing.

The screenshot shows the 'Import Data - Select the cells to import.' dialog box. The 'Sheet' is set to 'Testing'. The 'Cell range' is 'A:D'. The 'Define header row' checkbox is checked, and the 'Header row' is set to '1'. The dialog displays a table with 8 rows of data for the 'Testing' sheet:

	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

Buttons at the bottom: Previous, Next, Cancel.

8. Ubah Berangin menjadi type 'binominal', dan simpan dengan nama DataCuaca_Testing.

Import Data - Format your columns.

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

Import Data - Where to store the data?

Where to store the data?

Local Repository (Local)

- data
- processes
 - DataCuaca_Training (10/19/21 10:56 PM - 2 kB)

Name DataCuaca_Testing

Location //Local Repository/DataCuaca_Testing

Previous Finish Cancel

9. Menampilkan hasil import file Tabel_Cuaca.xls pada sheet2(Testing).

The screenshot shows the RapidMiner Studio interface with the Results view selected. The main window displays a table with 7 rows of weather data. The table has columns: Row No., Cuaca, Suhu, Kelembaban..., and Berangin. The data is as follows:

Row No.	Cuaca	Suhu	Kelembaban...	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

The interface also shows a Repository panel on the right with 'DataCuaca_Testing' and 'DataCuaca_Training' listed under the Local Repository. The bottom status bar indicates 'ExampleSet (7 examples, 0 special attributes, 4 regular attributes)'.

10. Membuat desain Naïve Bayes dengan mendrag DataCuaca_Training dan DataCuaca_Testing serta memasukkan operator Naïve Bayes dan Apply Model kedalam jendela process view.

The screenshot shows the RapidMiner Studio interface with the Process view selected. The workflow is as follows:

- Two 'Retrieve DataCuaca...' operators are connected to the 'Naïve Bayes' operator.
- The 'Naïve Bayes' operator is connected to the 'Apply Model' operator.
- The 'Apply Model' operator is connected to the 'res' output.

The interface also shows a Repository panel on the left with 'DataCuaca_Testing' and 'DataCuaca_Training' listed under the Local Repository. The bottom status bar indicates 'No results were found.' and 'Activate Wisdom of Crowds'.

11. Menjalankan proses Naïve Bayes dengan klik tombol panah biru/Run. Nilai kelas pada variable Y (Bermain_Tenis) rerata nilai confidence sebesar 0,856 untuk nilai TIDAK, dan 0,993 untuk nilai YA.

RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

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

Result History ExampleSet (Apply Model)

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)

Local Repository (Local)

Connections

data

processes

DataCuaca_Testing : 10/19/21 11:03 PM

DataCuaca_Training : 10/19/21 10:58 PM

DB (Legacy)

RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

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

Result History ExampleSet (Apply Model)

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

Name	Type	Missing	Statistics	Filter (7 / 7 attributes):
Prediction	Binominal	0	Negative TIDAK Positive YA	Values YA (5), TIDAK (2)
confidence(TIDAK)	Real	0	Min 0.007 Max 0.856	Average 0.353
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 (4), Mendung (2)
Suhu	Integer	0	Min 64 Max 83	Average 72.143
Kelembaban_udara	Integer	0	Min 65 Max 96	Average 79.571
Berangin	Binominal	0	Negative TIDAK Positive YA	Values YA (4), TIDAK (3)

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

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

Local Repository (Local)

Connections

data

processes

DataCuaca_Testing : 10/19/21 11:03 PM

DataCuaca_Training : 10/19/21 10:58 PM

DB (Legacy)

The screenshot displays two windows from the RapidMiner software suite. The main window on the left is RapidMiner Studio, showing a workflow with two 'ExampleSet' nodes connected to an 'Apply Model' node. The 'Result History' pane shows the output of the 'Apply Model' node, which is a table of 7 examples. The table has columns for Row No., prediction, confidence, and the original attributes: Cuaca and Suhu. The predictions are 'YA' for rows 1-3 and 5-6, and 'TIDAK' for rows 4 and 7. The confidences are 0.154, 0.498, 0.856, 0.019, 0.007, 0.371, and 0.568 respectively. The original attributes are Cerah, Cerah, Cerah, Mendung, Mendung, Hujan, and Hujan. The Suhu values are 75, 80, 83, 70, 68, 65, and 64. The bottom status bar indicates 'ExampleSet (7 examples, 3 special attributes, 4 regular attributes)'.

The window on the right is ARFF-Viewer, displaying the ARFF file 'HasilPrediksi.arff'. It shows the relation 'CuacaTesting_predicted' with 7 attributes: No, Nominal, Numeric, Numeric, Nominal, Numeric, and Nominal. The data is as follows:

No	1	2	3	4	5	6	7
Cerah	75.0	65.0	TIDAK	0.762765	YA		
Cerah	80.0	68.0	YA	0.097878	YA		
Cerah	83.0	87.0	YA	-0.676856	TIDAK		
Mend...	70.0	96.0	TIDAK	0.628523	YA		
Mend...	68.0	81.0	TIDAK	0.833996	YA		
Hujan	65.0	75.0	YA	0.253733	YA		
Hujan	64.0	85.0	YA	-0.160143	TIDAK		

1. Membuat file xls dan arff sesuai modul.

[illegible]

2. Membuat file arff Testing dan Training.

Tugas_Testing - Notepad

```
File Edit Format View Help
@relation Universitas

@attribute Jurusan_SMA {IPA, IPS, LAIN}
@attribute Gender {PRIA, WANITA}
@attribute Asal_Sekolah {SURAKARTA, LUAR}
@attribute Rerata_SKS real
@attribute Asisten {YA, TIDAK}
@attribute Lama_Studi {TEPAT, TERLAMBAT}

@data
LAIN,WANITA,SURAKARTA,18,TIDAK,?
IPA,PRIA,SURAKARTA,19,YA,?
LAIN,PRIA,SURAKARTA,19,TIDAK,?
IPS,PRIA,LUAR,17,TIDAK,?
LAIN,WANITA,SURAKARTA,17,TIDAK,?
IPA,WANITA,LUAR,18,YA,?
IPA,PRIA,SURAKARTA,18,TIDAK,?
IPA,PRIA,SURAKARTA,19,TIDAK,?
IPS,PRIA,LUAR,18,TIDAK,?
LAIN,WANITA,SURAKARTA,18,TIDAK,?
```

Tugas_Training - Notepad

```
File Edit Format View Help
@relation Universitas

@attribute Jurusan_SMA {IPA, IPS, LAIN}
@attribute Gender {PRIA, WANITA}
@attribute Asal_Sekolah {SURAKARTA, LUAR}
@attribute Rerata_SKS real
@attribute Asisten {YA, TIDAK}
@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
```

3. Setelah itu buka Weka Explorer, open file data Training, pilih tab Classify, set ke data Testing lalu klik Start.

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)

23:34:42 - bayes.NaiveBayes

Classifier output

```
[total] 15.0 9.0

Time taken to build model: 0 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Total Number of Instances 0
Ignored Class Unknown Instances 10

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC  ROC Area  PRC Area  Cla
          ?      ?      ?      ?      ?      ?      ?      ?      ?
Weighted Avg.  ?      ?      ?      ?      ?      ?      ?      ?      ?

=== Confusion Matrix ===

a b  <-- classified as
0 0 | a = TEPAT
0 0 | b = TERLAMBAT
```

Status

OK Log x0

5. Setelah itu save dengan nama HasilPrediksiTugas.arff, berikut adalah hasil prediksi Weka.

ARFF-Viewer - C:\Kuliah\SMT 5\Prak DWDM (Data Warehousing)\Modul 4\HasilPrediksiTugas.arff

File Edit View

HasilPrediksiTugas.arff

Relation: Universitas_predicted

No.	1: Jurusan_SMA	2: Gender	3: Asal_Sekolah	4: Rerata_SKS	5: Asisten	6: prediction margin	7: predicted Lama_Studi	8: Lama_Studi
	Nominal	Nominal	Nominal	Numeric	Nominal	Numeric	Nominal	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	

6. Lakukan langkah seperti di awal tadi untuk mendapatkan prediksi RapidMiner, Hasil prediksi data testing (Excel) menggunakan RapidMiner sebagai berikut.

<new process> - RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/Tabel_TugasTraining) ExampleSet (/Local Repository/Tabel_TugasTrain) ExampleSet (/Local Repository/Tabel_TugasTesting)

Result History ExampleSet (Apply Model)

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

Row No.	prediction(L...	confidence(L...	confidence(L...	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

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

Local Repository (Local)

Connections

data

processes

DataCuaca_Testing (10/19/21 11:03 PM - 2 kB)

DataCuaca_Training (10/19/21 10:56 PM - 2 kB)

Tabel_TugasTesting (10/20/21 12:02 AM - 2 kB)

Tabel_TugasTraining (10/20/21 12:01 AM - 3 kB)

DB (Legacy)

<new process> - RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

ExampleSet (/Local Repository/Tabel_TugasTraining) ExampleSet (/Local Repository/Tabel_TugasTrain) ExampleSet (/Local Repository/Tabel_TugasTesting)

Result History ExampleSet (Apply Model)

Name Type Missing Statist... Filter (8 / 8 attributes): Search for Attributes

✓ Prediction	prediction(Lama_Studi)	Binominal	0	Negative	Positive	Value
				TERLAMBAT	TEPAT	TEP
✓ Confidence_TERLAMBAT	confidence(TERLAMBAT)	Real	0	Min	Max	Average
				0.005	0.868	0.52
✓ Confidence_TEPAT	confidence(TEPAT)	Real	0	Min	Max	Average
				0.132	0.995	0.47
✓ Jurusan_SMA	Jurusan_SMA	Polynomial	0	Least	Most	Value
				IPS (2)	IPA (4)	IPA
✓ Gender	Gender	Polynomial	0	Least	Most	Value
				WANITA (4)	PRIA (6)	PRI
✓ Asal_Sekolah	Asal_Sekolah	Polynomial	0	Least	Most	Value
				LUAR (3)	SURAKARTA (7)	SUR
✓ Rerata_SKS	Rerata_SKS	Integer	0	Min	Max	Average
				17	19	18.0
✓ Asisten	Asisten	Binominal	0	Negative	Positive	Value
				TIDAK	YA	TID

Showing attributes 1 - 8 Examples: 10 Special Attributes: 3 Regular Attributes: 5

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

Local Repository (Local)

Connections

data

processes

DataCuaca_Testing (10/19/21 11:03 PM - 2 kB)

DataCuaca_Training (10/19/21 10:56 PM - 2 kB)

Tabel_TugasTesting (10/20/21 12:02 AM - 2 kB)

Tabel_TugasTraining (10/20/21 12:01 AM - 3 kB)

DB (Legacy)

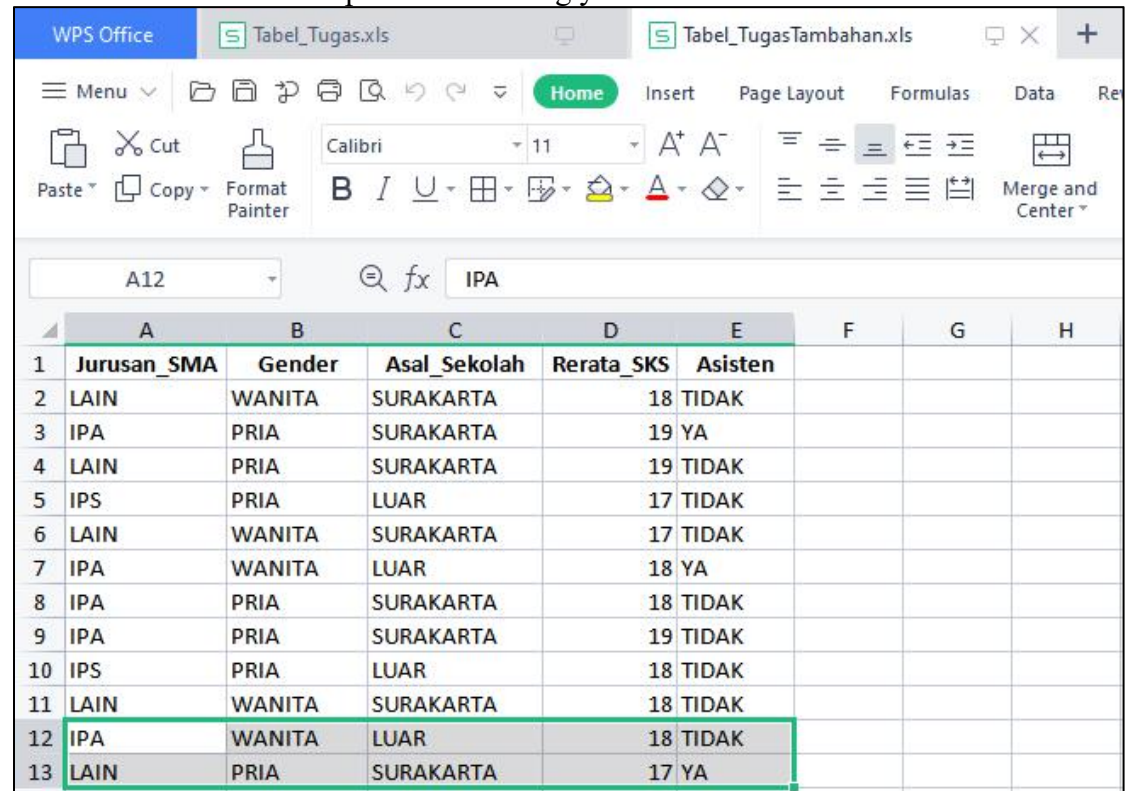
7. Nilai Rerata Confidance untuk atribut Lama_Studi dengan nilai TEPAT dan TERLAMBAT

- a. Nilai TEPAT : 0,995
- b. Nilai TERLAMBAT : 0,868

8. Jumlah orang yang akan lulus TEPAT dan TERLAMBAT:

- a. Tepat : 3 orang
- b. Terlambat : 7 orang

9. Menambahkan kondisi pada data Testing yaitu Dewi dan Jono.



	A	B	C	D	E	F	G	H
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			
12	IPA	WANITA	LUAR	18	TIDAK			
13	LAIN	PRIA	SURAKARTA	17	YA			

10. Hapus Testing dengan data kondisi baru.

The screenshot shows the RapidMiner Studio interface. The 'Process' tab is active, displaying a workflow diagram. The workflow starts with 'Retrieve Tabel_Tugas...' (highlighted in orange), followed by 'Naive Bayes', and then 'Apply Model'. The 'Retrieve Tabel_Tugas...' operator is connected to the 'Naive Bayes' operator, which is then connected to the 'Apply Model' operator. The 'Apply Model' operator is connected to a 'res' output. The 'Repository' panel on the left shows the 'data' folder containing 'Tabel_TugasTambahan' and 'Tabel_TugasTraining'. The 'Operators' panel on the left shows the 'Naive Bayes' operator selected. The 'Parameters' panel on the right shows the 'Retrieve Tabel_TugasTambahan (Retrieve)' operator with its repository entry set to 'epository/Tabel_TugasTambahan'.

11. Prediksi ketepatan lama studi si Dewi : TEPAT

The screenshot shows the 'Result History' table in RapidMiner Studio. The table displays 12 examples, with the 11th example highlighted in red. The 11th example is a correct prediction (TEPAT) with a confidence of 0.298. The table columns are: Row No., prediction(L..., confidence(L..., confidence(..., Jurusan_SMA, Gender, Asal_Sekolah, Rerata_SKS, and Asisten. The 11th example has a prediction of 'TEPAT', a confidence of 0.298, and a confidence of 0.702. The 12th example has a prediction of 'TEPAT', a confidence of 0.076, and a confidence of 0.924.

Row No.	prediction(L...	confidence(L...	confidence(...	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
11	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
12	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

12. Prediksi ketepatan lama studi si Jono : TEPAT

RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

ExampleSet (//Local Repository/Tabel_TugasTraining) ExampleSet (//Local Repository/Tabel_TugasTrain) ExampleSet (//Local Repository/Tabel_TugasTambahan) ExampleSet (//Local Repository/Tabel_TugasTesting) ExampleSet (Apply Model)

Result History

Open in Turbo Prep Auto Model Filter (12 / 12 examples): all

Row No.	prediction(L...	confidence(L...	confidence(L...	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
11	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
12	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA

ExampleSet (12 examples, 3 special attributes, 5 regular attributes)

Repository

- Training Resources (connected)
- Samples
- Community Samples (connected)
- Local Repository (Local)
- Connections
- data
- processes
 - DataCuaca_Testing (10/19/21 11:03 PM - 2 KB)
 - DataCuaca_Training (10/19/21 10:56 PM - 2 KB)
 - Tabel_TugasTambahan (10/20/21 12:18 AM - 2 KB)
 - Tabel_TugasTesting (10/20/21 12:02 AM - 2 KB)
 - Tabel_TugasTraining (10/20/21 12:01 AM - 3 KB)
- DB (Legacy)

13. Statistics

RapidMiner Studio Educational 9.10.000 @ DESKTOP-4Q38BK1

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Find data, operators, etc. All Studio

ExampleSet (//Local Repository/Tabel_TugasTraining) ExampleSet (//Local Repository/Tabel_TugasTrain) ExampleSet (//Local Repository/Tabel_TugasTambahan) ExampleSet (//Local Repository/Tabel_TugasTesting) ExampleSet (Apply Model)

Result History

Name Type Missing Statist... Filter (8 / 8 attributes): Search for Attributes

Name	Type	Missing	Statist...
Asisten	Binominal	0	Negative: TERLAMBAT (7), TEPAT (5)
Asas_Sekolah	Real	0	Min: 0.005, Max: 0.868, Average: 0.468
Asisten	Real	0	Min: 0.132, Max: 0.995, Average: 0.532
Asas_Sekolah	Polynomial	0	Least: IPS (2), Most: IPA (5), Values: IPA (5), LAIN (5), ... [1 more]
Asisten	Polynomial	0	Least: WANITA (5), Most: PRIA (7), Values: PRIA (7), WANITA (5)
Asas_Sekolah	Polynomial	0	Least: LUAR (4), Most: SURAKARTA (8), Values: SURAKARTA (8), LUAR (4)
Asisten	Integer	0	Min: 17, Max: 19, Average: 18

Showing attributes 1 - 8 Examples: 12 Special Attributes: 3 Regular Attributes: 5

Repository

- Training Resources (connected)
- Samples
- Community Samples (connected)
- Local Repository (Local)
- Connections
- data
- processes
 - DataCuaca_Testing (10/19/21 11:03 PM - 2 KB)
 - DataCuaca_Training (10/19/21 10:56 PM - 2 KB)
 - Tabel_TugasTambahan (10/20/21 12:18 AM - 2 KB)
 - Tabel_TugasTesting (10/20/21 12:02 AM - 2 KB)
 - Tabel_TugasTraining (10/20/21 12:01 AM - 3 KB)
- DB (Legacy)