

Nama : Roni Ardianzah
Nim : L200170073
Kelas : C

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

Menyiapkan File dengan format (EXCEL) dan (ARFF)

The screenshot shows two windows side-by-side. The left window is a text editor displaying an ARFF file named 'Tabel_Tugas.arff'. The right window is an Excel spreadsheet named 'Tabel_Tugas - Excel' showing the same data in a tabular format.

ARFF File Content:

```
@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,?
```

Excel Spreadsheet Data:

Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
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

Prediksi dengan menggunakan WEKA

The screenshot shows the WEKA interface. The 'Current relation' is 'Mahasiswa' with 6 attributes and 20 instances. The 'Selected attribute' is 'Jurusan_SMA' (Nominal, 3 distinct values). The 'Class' is 'Lama_Studi (Nom)'. A bar chart is displayed at the bottom right, showing the distribution of the 'Lama_Studi' class across the 'Jurusan_SMA' categories.

Selected attribute details:

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

Bar Chart Data:

Jurusan_SMA	Lama_Studi (Nom)	Count
IPA	0	10
IPA	1	10
IPS	0	6
IPS	1	6
LAIN	0	4
LAIN	1	4

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose ZeroR

Test options

☐ Use training set

☒ Supplied test set

☐ Cross-validation Folds 10

☐ Percentage split % 66

(Nom) Lama_Studi

Result list (right-click for options)

14:02:42 - rules.ZeroR

Classifier output

```

=== Classifier model (full training set) ===
ZeroR predicts class value: TEPAT
Time taken to build model: 0 seconds

=== Evaluation on test set ===
Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===
Total Number of Instances      0

=== Detailed Accuracy By Class ===

```

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

```

=== Confusion Matrix ===
a b  <-- classified as
0 0 | a = TEPAT
0 0 | b = TERLAMBAT

```

Status

OK

x0

Relation: Mahasiswa_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	

Pediksi dengan Rapidminer

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

Training

Import Data - Select the cells to import.

Select the cells to import.

Sheet: Training 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

Previous Next Cancel

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten	Lama_Studi
	polynomial	polynomial	polynomial	integer	polynomial	binomial label
1	IPS	WANITA	SURAKARTA	18	TIDAK	TERLAMBAT
2	IPA	PRIA	SURAKARTA	19	YA	TEPAT
3	LAIN	PRIA	SURAKARTA	19	TIDAK	TERLAMBAT
4	IPA	PRIA	LUAR	17	TIDAK	TERLAMBAT
5	IPA	WANITA	SURAKARTA	17	TIDAK	TEPAT
6	IPA	WANITA	LUAR	18	YA	TEPAT
7	IPA	PRIA	SURAKARTA	18	TIDAK	TERLAMBAT
8	IPA	PRIA	SURAKARTA	19	TIDAK	TEPAT
9	IPS	PRIA	LUAR	18	TIDAK	TERLAMBAT
10	LAIN	WANITA	SURAKARTA	18	TIDAK	TEPAT
11	IPA	WANITA	SURAKARTA	19	TIDAK	TEPAT
12	IPS	PRIA	SURAKARTA	20	TIDAK	TEPAT
13	IPS	PRIA	SURAKARTA	19	TIDAK	TEPAT

no problems.

Previous Next Cancel

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

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_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 ⓘ

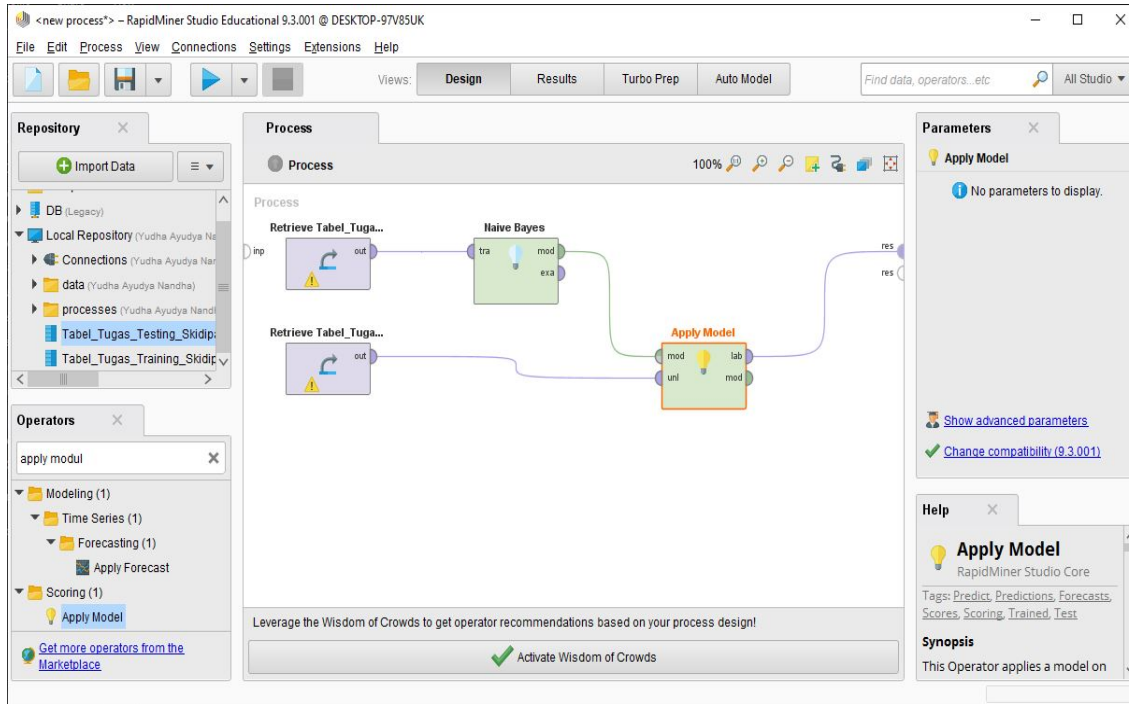
	Jurusan_SMA <i>polynomial</i>	Gender <i>polynomial</i>	Asal_Sekolah <i>polynomial</i>	Rerata_SKS <i>integer</i>	Asisten <i>binominal</i>
1	LAIN	WANITA	SURAKARTA	18	TIDAK
2	IPA	PRIA	SURAKARTA	19	YA
3	LAIN	PRIA	SURAKARTA	19	TIDAK
4	IPS	PRIA	LUAR	17	TIDAK
5	LAIN	WANITA	SURAKARTA	17	TIDAK
6	IPA	WANITA	LUAR	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

no problems.

Previous **Next** **Cancel**

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING



Result History

ExampleSet (Apply Model)

Open in: Turbo Prep Auto Model

Filter (10 / 10 examples): all

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

4. Nilai Rerata hasil dari prediksi RapidMiner untuk kolom Lama_Studi

Rerata untuk **Tepat** : 0,476

Rerata untuk **Tepat** : 0,524

MODUL 08

PRAKTIKUM DATA WAREHOUSING DAN DATA MINING

5. Dari hasil prediksi RapidMiner dapat disimpulkan bahwa 3 Mahasiswa akan lulus TEPAT dan 7 Mahasiswa akan lulus TERLAMBAT

6. Dewi Row No. 1

7. Jono Row No. 2

Row No.	prediction(L...	confidence{...	confidence{...	Jurusan_SMA	Gender	Asal_Sekolah	Rerata_SKS	Asisten
1	TEPAT	0.298	0.702	IPA	WANITA	LUAR	18	TIDAK
2	TEPAT	0.076	0.924	LAIN	PRIA	SURAKARTA	17	YA