

## Praktikum Data Warehouse dan Data Mining

Nama : Aldias Ibnu Habib

NIM : L200170166

Kelas : F

Modul : 9

The image displays two screenshots from a Windows desktop environment. The top screenshot shows the Weka Explorer application window. The 'Classify' tab is active, and the 'J48 -C 0.25 -M 2' classifier is selected. The 'Test options' section shows 'Use training set' selected. The 'Classifier output' pane displays the following text:

```
Test mode: evaluate on training data

=== Classifier model (full training set) ===

J48 pruned tree

Cuaca = Cerah
| Kelembaban_Udara <= 75: Ya (2.0)
| Kelembaban_Udara > 75: Tidak (3.0)
Cuaca = Mendung: Ya (4.0)
Cuaca = Hujan
| Berangin = Ya: Tidak (2.0)
| Berangin = Tidak: Ya (3.0)

Number of Leaves : 5
Size of the tree : 8

Time taken to build model: 0.02 seconds

=== Evaluation on training set ===
```

The bottom screenshot shows the Weka Classifier Tree Visualizer window, titled 'Weka Classifier Tree Visualizer: 12:46:50 - trees.J48 (Cuaca)'. It displays a visual representation of the decision tree structure:

```
graph TD
    Cuaca -- "= Cerah" --> Kelembaban_Udara
    Cuaca -- "= Mendung" --> Ya40[Ya (4.0)]
    Cuaca -- "= Hujan" --> Berangin
    Kelembaban_Udara -- "<= 75" --> Ya20[Ya (2.0)]
    Kelembaban_Udara -- "> 75" --> Tidak30[Tidak (3.0)]
    Berangin -- "= Ya" --> Tidak20[Tidak (2.0)]
    Berangin -- "= Tidak" --> Ya30[Ya (3.0)]
```

The visualizer also shows the same text output as the Weka Explorer window, including the model details and evaluation results.

**RapidMiner Studio Trial 9.3.001 @ LABSI-15-PC**

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

**Repository**

- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-15)
  - connections (LABSI-15)
  - data (LABSI-15)
  - processes (LABSI-15)
    - Tabel\_CuacaTesting (LABSI-15 - v1)
    - Tabel\_CuacaTraining (LABSI-15 - v1)

**Operators**

performance

- Modeling (1)
- Validation (19)
  - Performance (17)
    - Predictive (7)
      - Performance (Classification)
      - Performance (Binominal Cla)
      - Performance (Regression)

We found "Model Management" in the Marketplace. [Show me!](#)

**Process**

Process

100%

Retrieve Tabel\_Cuac... Cross Validation

Process Parameters

- logverbosity: init
- logfile:
- resultfile:
- random seed: 2001
- send mail: never
- encoding: SYSTEM

[Hide 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**

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

Activate Wisdom of Crowds

**RapidMiner Studio Trial 9.3.001 @ LABSI-15-PC**

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

**Repository**

- Community Samples (connected)
- DB (Legacy)
- Local Repository (LABSI-15)
  - connections (LABSI-15)
  - data (LABSI-15)
  - processes (LABSI-15)
    - Tabel\_CuacaTesting (LABSI-15 - v1)
    - Tabel\_CuacaTraining (LABSI-15 - v1)

**Operators**

performance

- Modeling (1)
- Validation (19)
  - Performance (17)
    - Predictive (7)
      - Performance (Classification)
      - Performance (Binominal Cla)
      - Performance (Regression)

We found "Model Management" in the Marketplace. [Show me!](#)

**Process**

Process > Cross Validation

100%

Training Testing

Decision Tree Apply Model Performance

Process Parameters

**Decision Tree**

- criterion: information...
- maximal depth: 10
- ☒ apply pruning
- confidence: 0.1
- ☒ apply prepruning
- minimal gain: 0.01

[Hide advanced parameters](#)

**Help**

**Decision Tree**

Concurrency

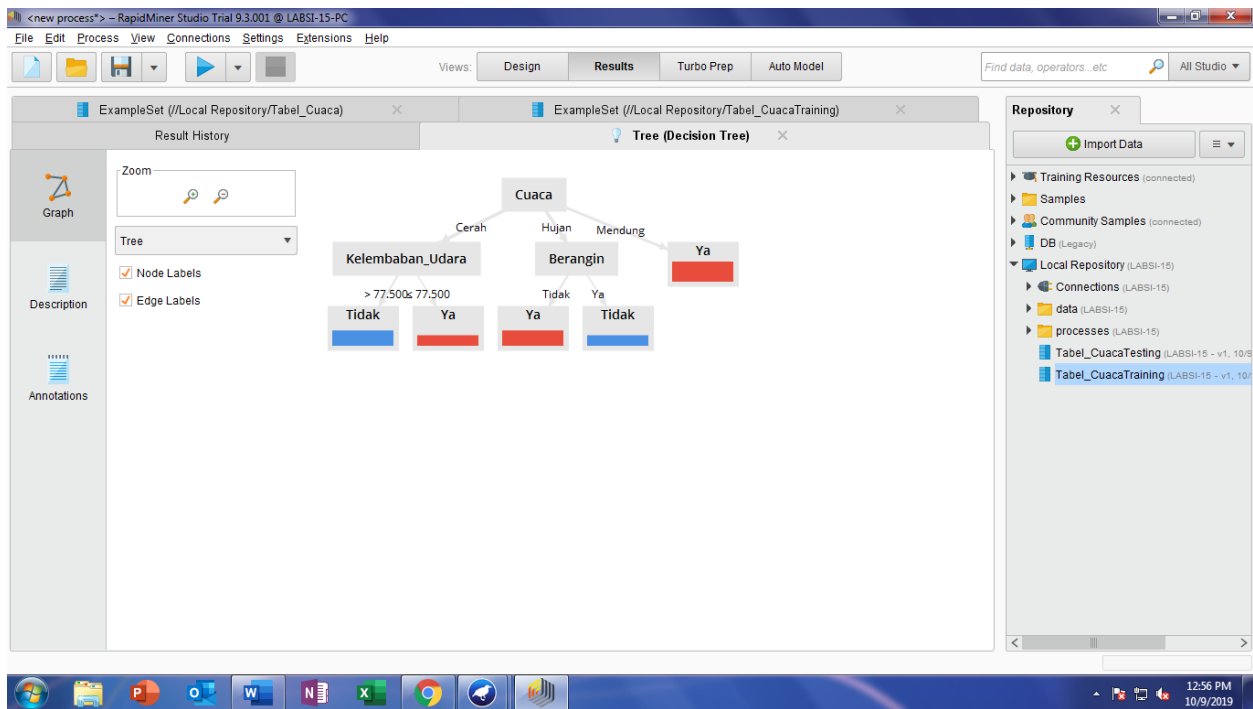
Tags: Supervised, Classification, Regression, Model, Trees

**Synopsis**

This Operator generates a decision tree model, which can be used for classification and regression.

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

Activate Wisdom of Crowds

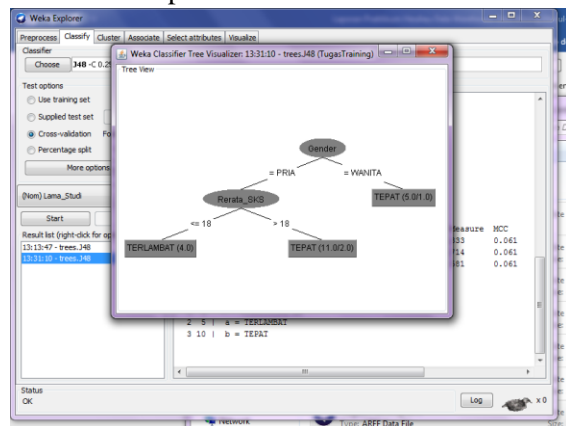


1.

Cuaca	Suhu	Kelembaban_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	Ya	Ya
Hujan	64	85	Ya	Tidak

2.

a. Pohon keputusan berdasarkan data berikut



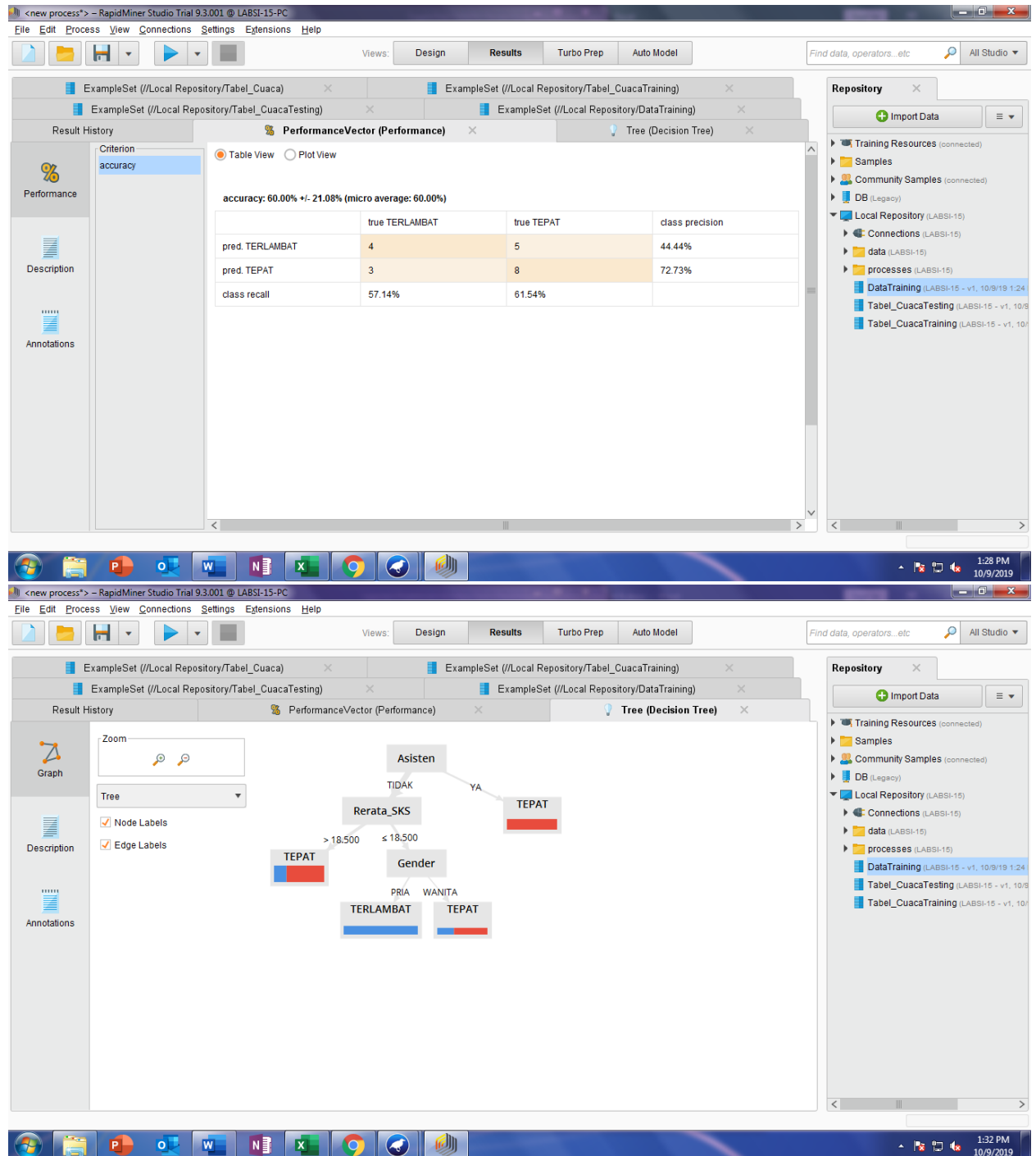
b.

i) Jumlah simpul daun pada pohon keputusan 3

- ii) Jumlah simpul keseluruhan pada pohon keputusan 5
- iii) Waktu yang dibutuhkan untuk proses pelatihan = 0,2 seconds
- iv) Tingkat ketepatan klasifikasi 85%
- v) Tingkat ketidaktepatan 15%

3. Data training file excel dari Tugas nomor 1 modul 6 seperti berikut.

a.



b.

- i) Jumlah simpul daun pada pohon keputusan 3
- ii) Jumlah simpul keseluruhan pada pohon keputusan 5
- iii) Waktu yang dibutuhkan untuk proses pelatihan = 0,2 seconds
- iv) Tingkat ketepatan klasifikasi 85%
- v) Tingkat ketidaktepatan 15%

Klasifikasi yang terbentuk yaitu:

- i. Orang yang akan Lama Studi Tepat jika kondisi sebagai berikut:  
Asisten=Tidak, Rerata\_SKS>18.500  
Asisten=Tidak, Rerata\_SLS<=18.500, Gender=Wanita  
Asisten=Ya
- ii. Orang akan Lama Studi Terlambat jika kondisi sebagai berikut:  
Asisten=Tidak, Rerata\_SKS<=18.500, Gender=Pria