

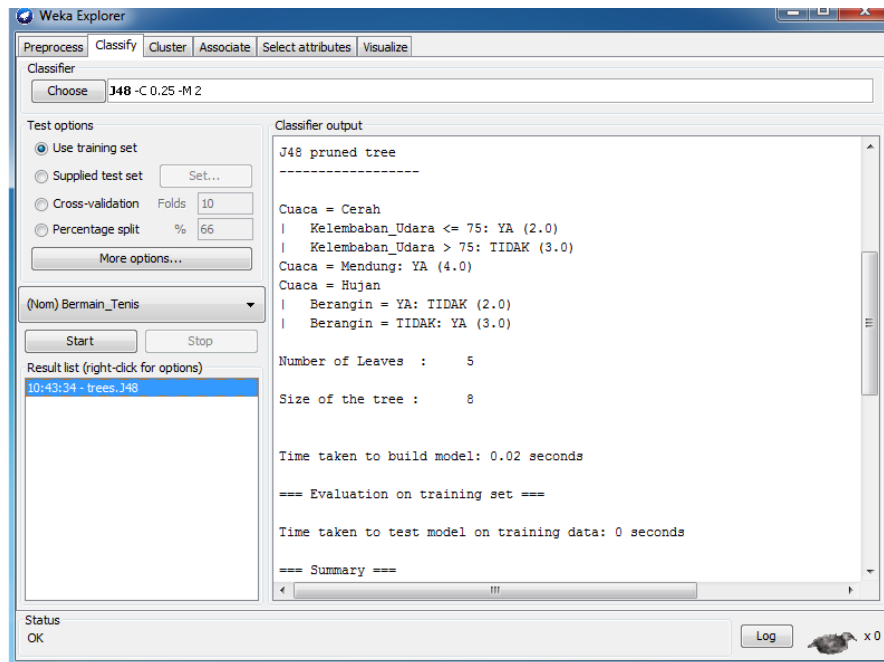
PRAKTIKUM DATA WAREHOUSING dan DATA MINING

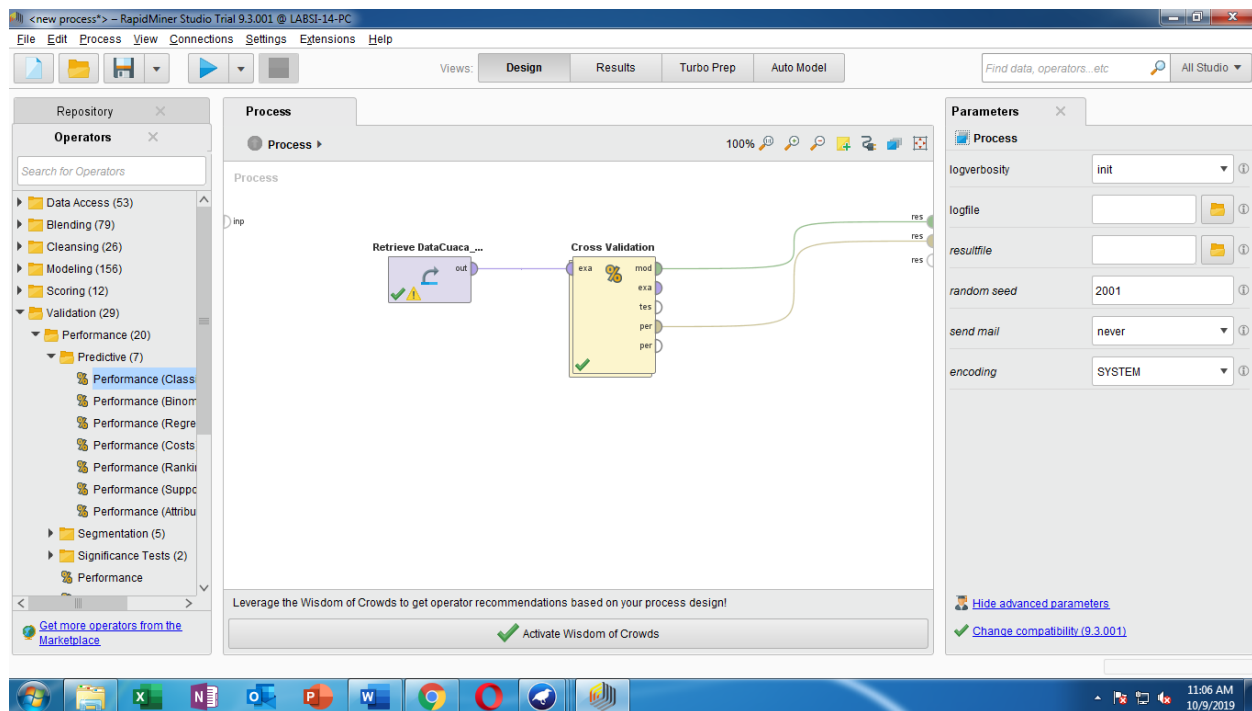
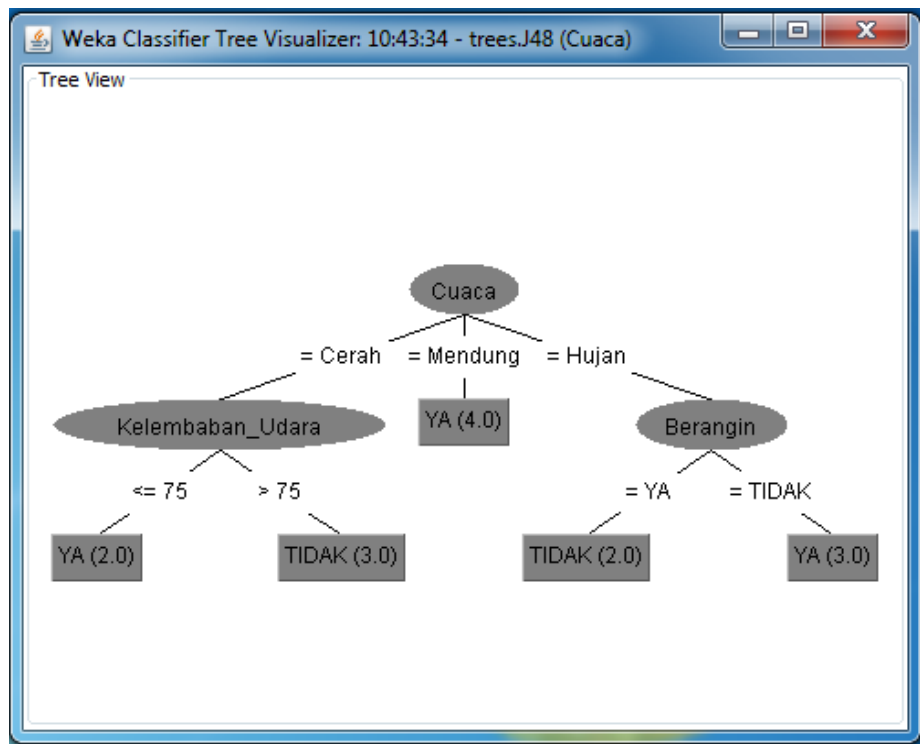
MODUL 9

NAMA : AL ADIAT RUSSETYA TAMORA

NIM : L200170037

KELAS: B





<new process*> - RapidMiner Studio Trial 9.3.001 @ LABSI-14-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

Repository

Operators

Search for Operators

- Data Access (53)
- Blending (79)
- Cleansing (26)
- Modeling (156)
- Scoring (12)
- Validation (29)
 - Performance (20)
 - Predictive (7)
 - Performance (Class)
 - Performance (Binom)
 - Performance (Regre)
 - Performance (Costs)
 - Performance (Ranki)
 - Performance (Supp)
 - Performance (Attribu)
 - Segmentation (5)
 - Significance Tests (2)
 - Performance

Get more operators from the Marketplace

Process

Process > Cross Validation

100%

Training

Decision Tree

Testing

Apply Model

Performance

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

Activate Wisdom of Crowds

Parameters

Cross Validation

- ☐ split on batch attribute
- ☐ leave one out
- number of folds: 10
- sampling type: automatic
- ☐ use local random seed
- ☒ enable parallel execution

Hide advanced parameters

Change compatibility (9.3.001)

Double-click to enter subprocess, drag to move.

<new process*> - RapidMiner Studio Trial 9.3.001 @ LABSI-14-PC

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model

Find data, operators, etc. All Studio

Result History

PerformanceVector (Performance) Tree (Decision Tree)

Criterion: accuracy

Table View Plot View

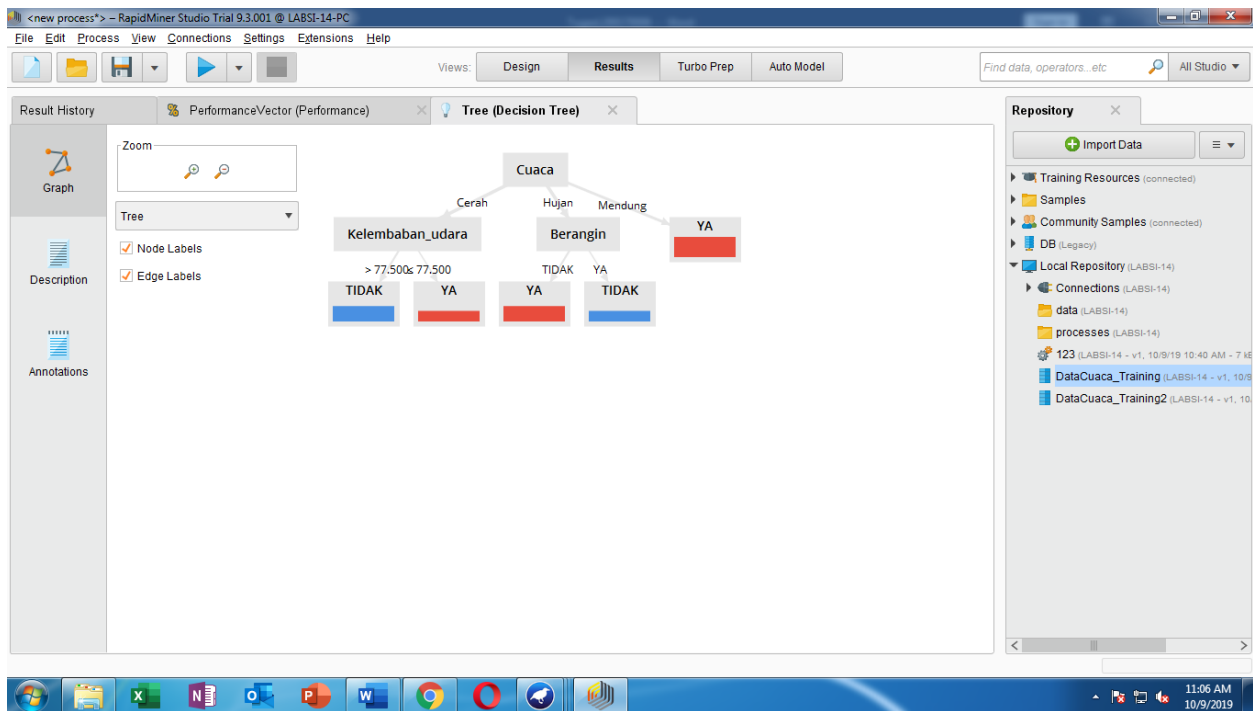
accuracy: 60.00% +/- 45.95% (micro average: 64.29%)

| | true TIDAK | true YA | class precision |
|--------------|------------|---------|-----------------|
| pred. TIDAK | 2 | 2 | 50.00% |
| pred. YA | 3 | 7 | 70.00% |
| class recall | 40.00% | 77.78% | |

Repository

Import Data

- Training Resources (connected)
 - Samples
 - Community Samples (connected)
 - DB (Legacy)
- Local Repository (LABSI-14)
 - Connections (LABSI-14)
 - data (LABSI-14)
 - processes (LABSI-14)
 - 123 (LABSI-14 - v1, 10/9/19 10:40 AM - 7 MB)
 - DataCuaca_Training (LABSI-14 - v1, 10/9/19 10:40 AM - 7 MB)
 - DataCuaca_Training2 (LABSI-14 - v1, 10/9/19 10:40 AM - 7 MB)



Tugas

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.

Weka Explorer

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

Classifier: Choose **J48 -C 0.25 -M 2**

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)

- 09:50:24 - trees.J48

Classifier output

Kappa statistic 0.6341
Mean absolute error 0.2436
Root mean squared error 0.349
Relative absolute error 53.0693 %
Root relative squared error 73.1456 %
Coverage of cases (0.95 level) 100 %
Mean rel. region size (0.95 level) 90 %
Total Number of Instances 20

=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC |
|---------------|---------|---------|-----------|--------|-----------|-------|
| | 0.571 | 0.000 | 1.000 | 0.571 | 0.727 | 0.681 |
| | 1.000 | 0.429 | 0.813 | 1.000 | 0.897 | 0.681 |
| Weighted Avg. | 0.850 | 0.279 | 0.878 | 0.850 | 0.837 | 0.681 |

=== Confusion Matrix ===

a b <-- classified as

| | | |
|---|----|---------------|
| 4 | 3 | a = TERLAMBAT |
| 0 | 13 | b = TEPAT |

Status: OK

Log x 0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose **J48 -C 0.25 -M 2**

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):

- 09:50:24 - trees.J48

Classifier output:

Time taken to test model on training data: 0 seconds

=== Summary ===

| | | | |
|------------------------------------|---------|----|---|
| Correctly Classified Instances | 17 | 85 | % |
| Incorrectly Classified Instances | 3 | 15 | % |
| Kappa statistic | 0.6341 | | |
| Mean absolute error | 0.2436 | | |
| Root mean squared error | 0.349 | | |
| Relative absolute error | 53.0693 | % | |
| Root relative squared error | 73.1456 | % | |
| Coverage of cases (0.95 level) | 100 | % | |
| Mean rel. region size (0.95 level) | 90 | % | |
| Total Number of Instances | 20 | | |

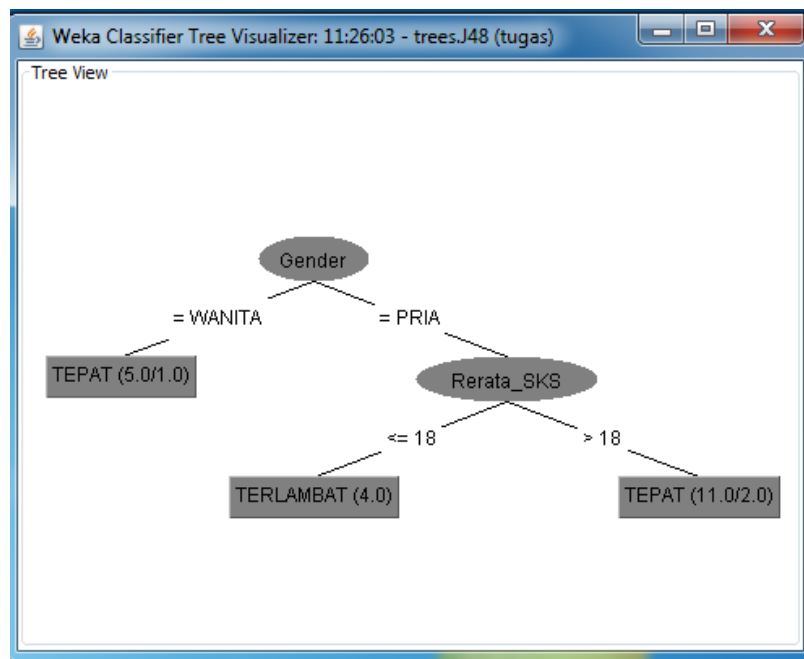
=== Detailed Accuracy By Class ===

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC |
|---------------|---------|---------|-----------|--------|-----------|-------|
| | 0.571 | 0.000 | 1.000 | 0.571 | 0.727 | 0.681 |
| | 1.000 | 0.429 | 0.813 | 1.000 | 0.897 | 0.681 |
| Weighted Avg. | 0.850 | 0.279 | 0.878 | 0.850 | 0.837 | 0.681 |

=== Confusion Matrix ===

Status: OK

Log x 0



Jumlah simpul daun pada pohon keputusan = 3

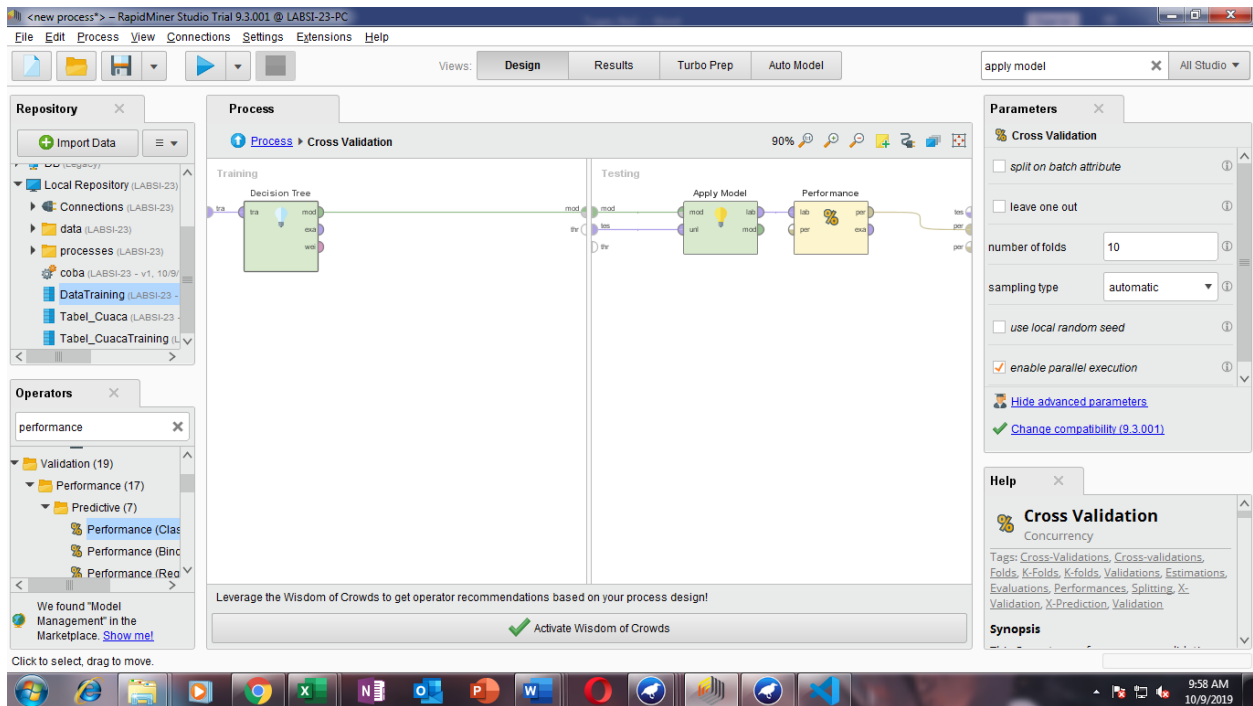
Jumlah simpul keseluruhan pada pohon keputusan = 5

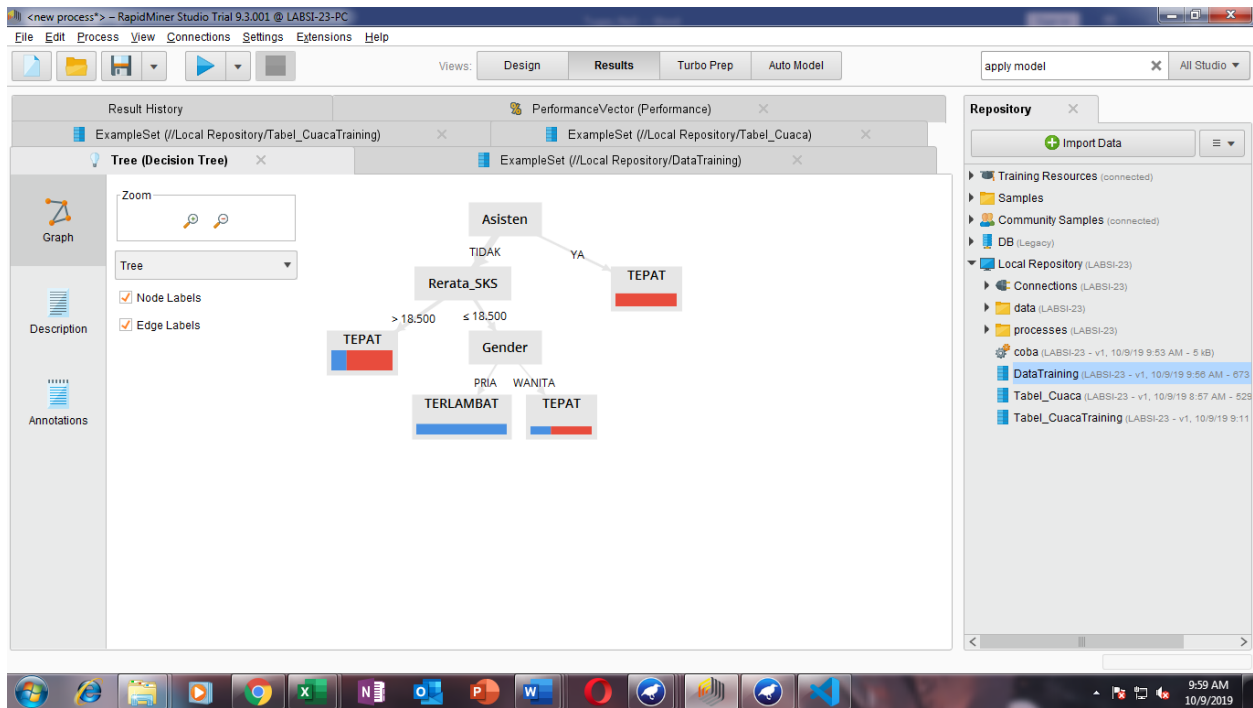
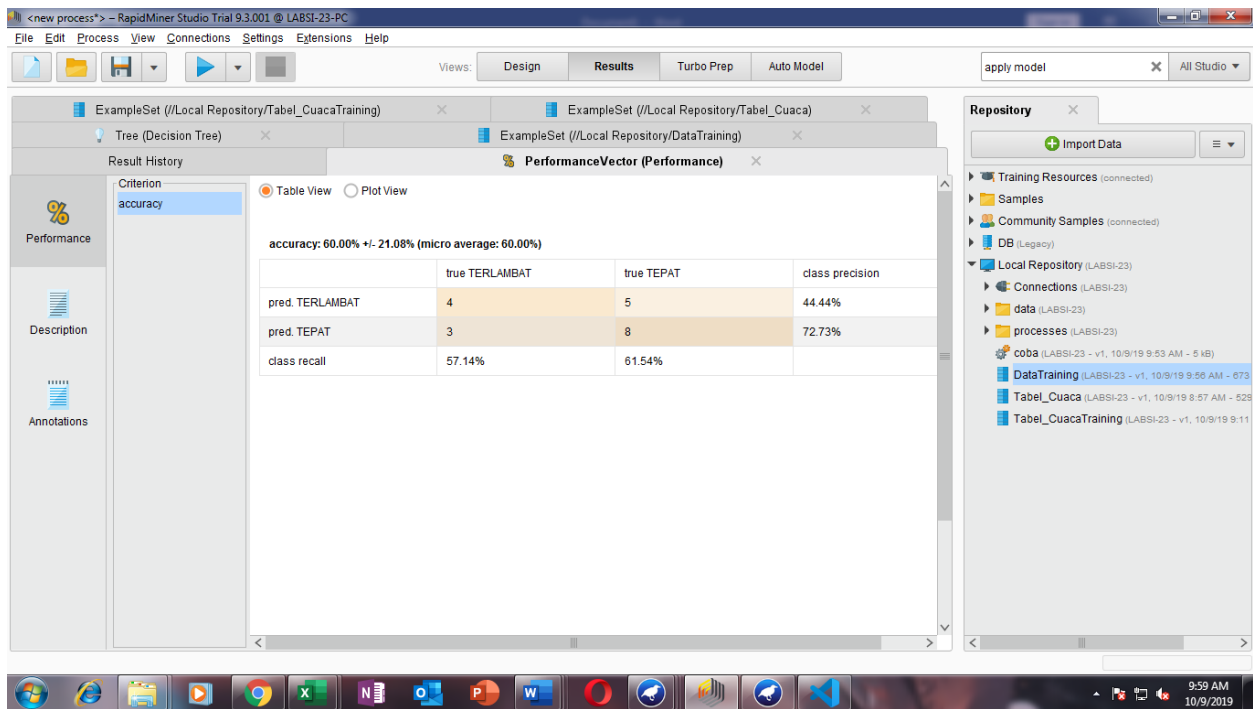
Waktu yang dibutuhkan untuk proses pelatihan = 0 detik

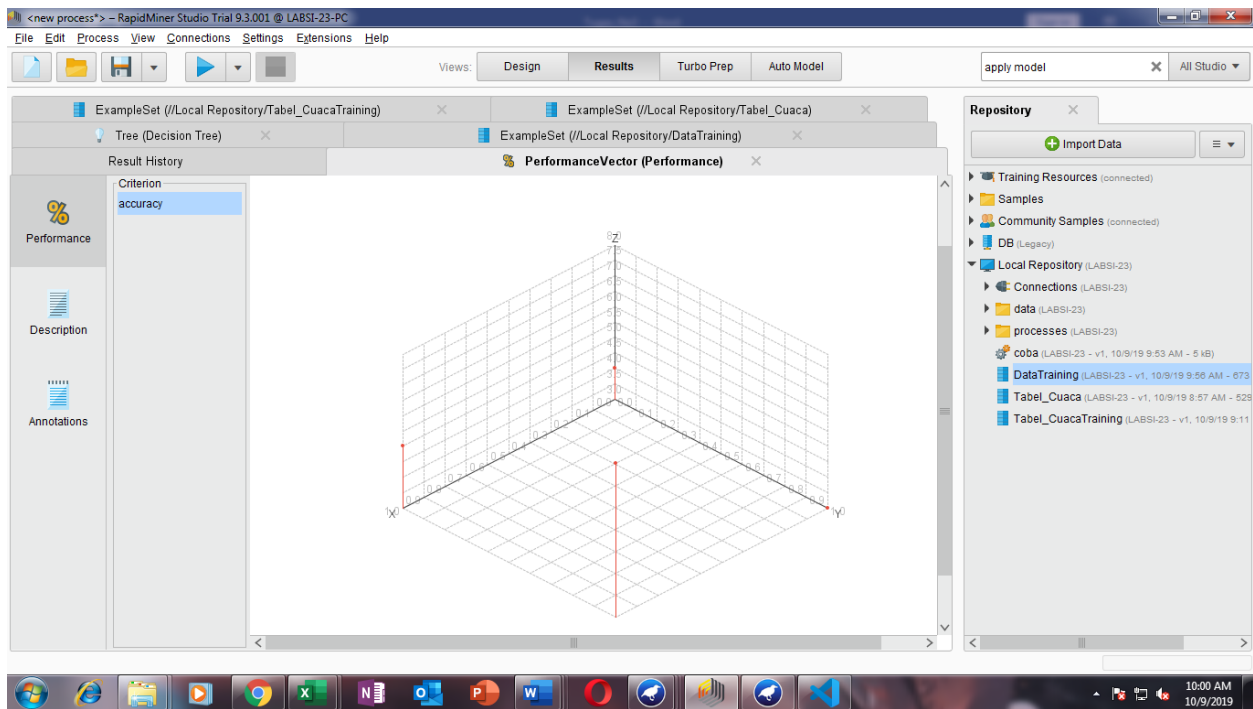
Tingkat ketepatan klasifikasi = 85 %

Tingkat ketidaktepatan klasifikasi = 15 %

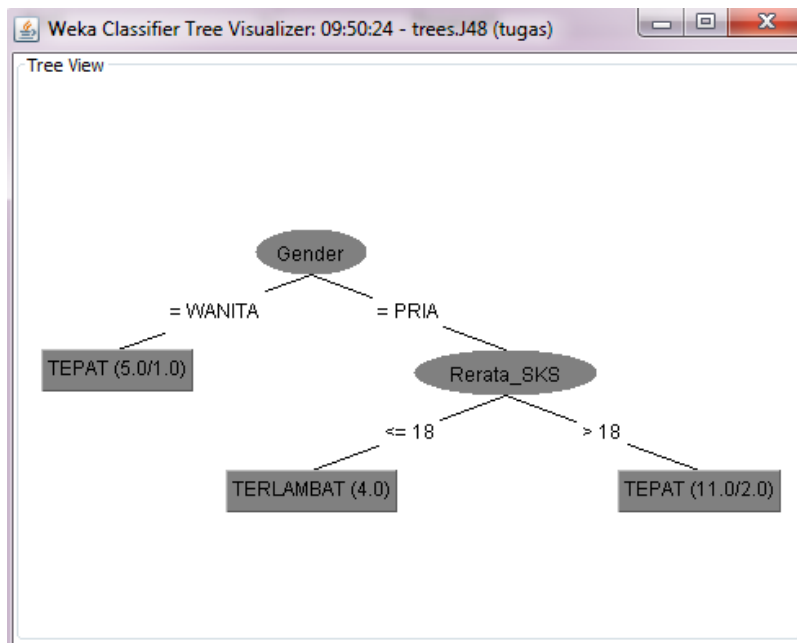
3.







4.



Klasifikasi yang terbentuk yaitu :

Seseorang akan Lama Studi (TEPAT) jika kondisi :

Gender = Wanita (Nilai atribut lain diabaikan)

Gender = Pria, Rerata_SKS > 18 (Nilai atribut lain diabaikan)

Seseorang akan Lama Studi (TERLAMBAT) jika kondisi :

Gender = Pria, Rerata_SKS \leq 18 (Nilai atribut lain diabaikan)