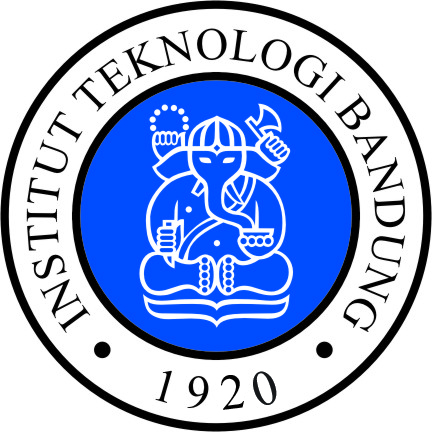
**Laporan Tugas Besar 1**

**IF4071 Pembelajaran Mesin**



Disusun oleh:

Felicia Christie – 13512039

Ivana Clairine Irsan – 13512041

Tegar Aji Pangestu – 13512061

Jurusan Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung

2015

1. Uji coba algoritma ID3 dengan myID3

* Dataset Weather (Nominal)

Untuk dataset ini, dilakukan beberapa macam uji, yaitu pengujian dengan training set sebagai test set, *10-Fold Cross Validation,* serta 50% *Percentage Split*. Setiap pengujian myID3 menghasilkan hasil yang sesuai dengan hasil dari algoritma Weka.

* + Uji dengan algoritma ID3 Weka, training set sebagai test set

Correctly Classified Instances 14 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 0 1 1 1 1 yes

1 0 1 1 1 1 no

Weighted Avg. 1 0 1 1 1 1

=== Confusion Matrix ===

a b <-- classified as

9 0 | a = yes

0 5 | b = no

* + Uji dengan algoritma myID3, training set sebagai test set

Correctly Classified Instances 14 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 0 1 1 1 1 yes

1 0 1 1 1 1 no

Weighted Avg. 1 0 1 1 1 1

=== Confusion Matrix ===

a b <-- classified as

9 0 | a = yes

0 5 | b = no

* + Uji dengan algoritma ID3 Weka, 10-Fold Cross Validation

Correctly Classified Instances 12 85.7143 %

Incorrectly Classified Instances 2 14.2857 %

Kappa statistic 0.6889

Mean absolute error 0.1429

Root mean squared error 0.378

Relative absolute error 30 %

Root relative squared error 76.6097 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.889 0.2 0.889 0.889 0.889 0.844 yes

0.8 0.111 0.8 0.8 0.8 0.844 no

Weighted Avg. 0.857 0.168 0.857 0.857 0.857 0.844

=== Confusion Matrix ===

a b <-- classified as

8 1 | a = yes

1 4 | b = no

* + Uji dengan algoritma myID3, 10-Fold Cross Validation

Correctly Classified Instances 12 85.7143 %

Incorrectly Classified Instances 2 14.2857 %

Kappa statistic 0.6889

Mean absolute error 0.1429

Root mean squared error 0.378

Relative absolute error 30 %

Root relative squared error 76.6097 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.889 0.2 0.889 0.889 0.889 0.844 yes

0.8 0.111 0.8 0.8 0.8 0.844 no

Weighted Avg. 0.857 0.168 0.857 0.857 0.857 0.844

=== Confusion Matrix ===

a b <-- classified as

8 1 | a = yes

1 4 | b = no

* + Uji dengan algoritma ID3 Weka untuk 50% *percentage split*

Correctly Classified Instances 5 71.4286 %

Incorrectly Classified Instances 2 28.5714 %

Kappa statistic 0.4615

Mean absolute error 0.2857

Root mean squared error 0.5345

Relative absolute error 60 %

Root relative squared error 111.6313 %

Total Number of Instances 7

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.6 0 1 0.6 0.75 0.8 yes

1 0.4 0.5 1 0.667 0.8 no

Weighted Avg. 0.714 0.114 0.857 0.714 0.726 0.8

=== Confusion Matrix ===

a b <-- classified as

3 2 | a = yes

0 2 | b = no

* + Uji dengan algoritma myID3 untuk 50% *percentage split*

Correctly Classified Instances 5 71.4286 %

Incorrectly Classified Instances 2 28.5714 %

Kappa statistic 0.4615

Mean absolute error 0.2857

Root mean squared error 0.5345

Relative absolute error 60 %

Root relative squared error 111.6313 %

Total Number of Instances 7

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.6 0 1 0.6 0.75 0.8 yes

1 0.4 0.5 1 0.667 0.8 no

Weighted Avg. 0.714 0.114 0.857 0.714 0.726 0.8

=== Confusion Matrix ===

a b <-- classified as

3 2 | a = yes

0 2 | b = no

1. Uji coba algoritma J48 dengan myJ48

* Dataset Weather Nominal

Untuk J48, terdapat banyak perbedaan diantara hasil J48 buatan kami dengan J48 dari Weka.

* + J48 Weka, 10Fold Cross-Validation

Correctly Classified Instances 7 50 %

Incorrectly Classified Instances 7 50 %

Kappa statistic -0.0426

Mean absolute error 0.4167

Root mean squared error 0.5984

Relative absolute error 87.5 %

Root relative squared error 121.2987 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.556 0.6 0.625 0.556 0.588 0.633 yes

0.4 0.444 0.333 0.4 0.364 0.633 no

Weighted Avg. 0.5 0.544 0.521 0.5 0.508 0.633

=== Confusion Matrix ===

a b <-- classified as

5 4 | a = yes

3 2 | b = no

* + myJ48, 10 Fold Cross-Validation

Correctly Classified Instances 12 85.7143 %

Incorrectly Classified Instances 2 14.2857 %

Kappa statistic 0.6889

Mean absolute error 0.1429

Root mean squared error 0.378

Relative absolute error 30 %

Root relative squared error 76.6097 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.889 0.2 0.889 0.889 0.889 0.844 yes

0.8 0.111 0.8 0.8 0.8 0.844 no

Weighted Avg. 0.857 0.168 0.857 0.857 0.857 0.844

=== Confusion Matrix ===

a b <-- classified as

8 1 | a = yes

1 4 | b = no

* + J48 Weka, percentage split 80%

Correctly Classified Instances 2 66.6667 %

Incorrectly Classified Instances 1 33.3333 %

Kappa statistic 0

Mean absolute error 0.4545

Root mean squared error 0.4724

Relative absolute error 98.4848 %

Root relative squared error 99.6187 %

Total Number of Instances 3

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 1 0.667 1 0.8 0.5 yes

0 0 0 0 0 0.5 no

Weighted Avg. 0.667 0.667 0.444 0.667 0.533 0.5

=== Confusion Matrix ===

a b <-- classified as

2 0 | a = yes

1 0 | b = no

* + myJ48, percentage split 80%

Correctly Classified Instances 3 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 3

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 0 1 1 1 1 yes

1 0 1 1 1 1 no

Weighted Avg. 1 0 1 1 1 1

=== Confusion Matrix ===

a b <-- classified as

2 0 | a = yes

0 1 | b = no

* + J48 Weka, full training set

Correctly Classified Instances 14 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 0 1 1 1 1 yes

1 0 1 1 1 1 no

Weighted Avg. 1 0 1 1 1 1

=== Confusion Matrix ===

a b <-- classified as

9 0 | a = yes

0 5 | b = no

* + myJ48, full training set

Correctly Classified Instances 14 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 14

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

1 0 1 1 1 1 yes

1 0 1 1 1 1 no

Weighted Avg. 1 0 1 1 1 1

=== Confusion Matrix ===

a b <-- classified as

9 0 | a = yes

0 5 | b = no