Introduction to ML - Decision Tree Coursework Report
COMP70050

Panagiotis Christodoulou
Evangelos Georgiadis
Matthew Murphy
Anjali Ramesh

Step 3: Evaluation
Cross validation classification metrics

Clean Dataset

Confusion matrix

	Room 1	Room 2	Room 3	Room 4
Room 1	49.30	0	0.30	0.50
Room 2	0	48.10	2.30	0
Room 3	0.10	1.90	47.20	0.10
Room 4	0.60	0.10	0.20	49.40

Accuracy: 0.970

Average Depth: 12.9

Precision

Room 1	0.984
Room 2	0.954
Room 3	0.957
Room 4	0.984

Recall

Room 1	0.986
Room 2	0.962
Room 3	0.944
Room 4	0.988

F1-Score

Room 1	0.985
Room 2	0.958
Room 3	0.951
Room 4	0.986

Noisy Dataset

Confusion matrix

	Room 1	Room 2	Room 3	Room 4
Room 1	39.00	3.10	3.10	4.30
Room 2	2.70	39.90	3.40	2.00
Room 3	3.40	4.70	41.50	3.40
Room 4	3.90	2.00	3.50	40.10

Accuracy: 0.803

Average Depth: 20.1

Precision

Room 1	0.788
Room 2	0.831
Room 3	0.783
Room 4	0.81

Recall

Room 1	0.796
Room 2	0.803
Room 3	0.806
Room 4	0.805

F1-Score

Room 1	0.792
Room 2	0.817
Room 3	0.794
Room 4	0.808

Results analysis

The confusion matrix was computed by averaging the confusion matrix of each k-fold (10 in total). After evaluating the tree using 10 fold cross validation on the clean dataset, Room 1 and Room 4 were consistently classified more accurately than Room 2 and Room 3, which can be

seen across the precision, recall, and F1-score metrics. For example, the precision for Room 1 and 4 are both 0.984, while for Rooms 2 and 3 they are closer to 0.95. The confusion matrix also shows evidence of the same, as Room 1 and 4 are classified correctly more often than Room 2 and 3. There is less of a clear difference in classification of rooms in the noisy dataset, as each metric shows a slightly different result. In some metrics such as precision and F1-score, Rooms 2 and 4 are more correctly recognized, but the difference between each room's metrics is less significant than in the clean data.

Dataset differences

There is a difference when using the decision tree on the clean and noisy datasets. The most noticeable difference comes from the accuracy metric, which shows that the decision tree correctly classified signals in the clean dataset 97.5% of the time, while it only correctly classified signals in the noisy dataset 80.5% of the time in the noisy dataset. This could be due to the tree overfitting because of random signals in the data that do not follow a pattern or because of increased variance in the data itself. The tree fit on the noisy dataset also has an average depth of 19.6 as opposed to the clean dataset's 13.5, showing overfitting.

Step 4: Pruning and Evaluation

Cross validation classification metrics

Clean Dataset

Confusion matrix

	Room 1	Room 2	Room 3	Room 4
Room 1	49.94	0	0.31	0.32
Room 2	0	47.95	1.13	0
Room 3	0.05	2.04	48.3	0.12
Room 4	0	0	0.25	49.55

Accuracy: 0.978

Average Depth: 8.8

Noisy Dataset

Confusion matrix

	Room 1	Room 2	Room 3	Room 4
Room 1	45.18	1.78	1.87	1.81
Room 2	0.91	45.35	2.4	1.2
Room 3	1.1	1.56	46.02	1.33
Room 4	1.8	0.98	1.2	45.45

Accuracy: 0.910

Average Depth: 14.5

Precision

Room 1	0.987
Room 2	0.977
Room 3	0.956
Room 4	0.995

Precision

Room 1	0.892
Room 2	0.910
Room 3	0.920
Room 4	0.919

Recall

Room 1	0.999
Room 2	0.959
Room 3	0.966
Room 4	0.991

Recall

Room 1	0.922
Room 2	0.913
Room 3	0.894
Room 4	0.913

F1-Score

Room 1	0.993
Room 2	0.968
Room 3	0.961
Room 4	0.993

F1-Score

Room 1	0.907
Room 2	0.911
Room 3	0.907
Room 4	0.916

Result analysis after pruning

Pruning notably enhanced the model's performance on both clean and noisy datasets. In the clean data, there was a subtle improvement, with accuracy rising from 0.97 to 0.979. However, the noisy data saw a significant accuracy boost from 0.803 to 0.91. Post-pruning, precision, recall, and F1 scores in both datasets also improved, reflecting a more accurate and robust model after removing less important decision paths or "branches." These improvements signify that pruning effectively minimized overfitting, enhancing the model's predictive quality.

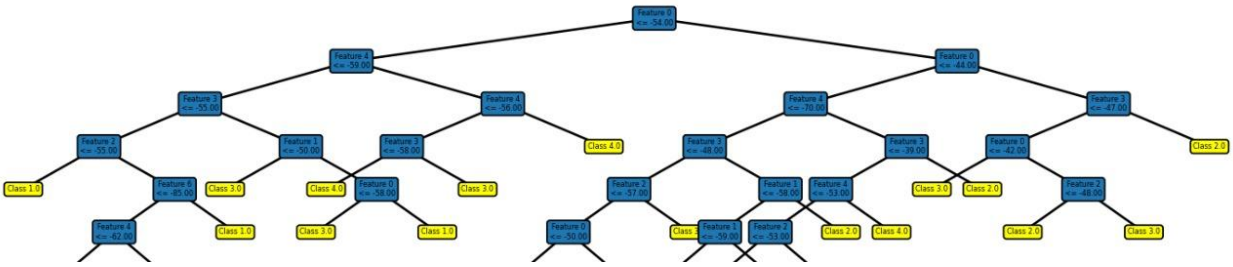
Depth analysis

Before pruning, the trees had substantial depths (13.5 for clean and 20.1 for noisy data), indicative of potential overfitting, especially in the noisy dataset. After pruning, the depths reduced significantly (8.8 for clean and 14.5 for noisy data), demonstrating simplification. A reduced tree depth post-pruning corresponded with improved prediction accuracies in both datasets, indicating a negative correlation between excessive depth and predictive accuracy.

Simplifying the trees through pruning, therefore, seemed instrumental in enhancing model performance by preventing overfitting and improving generalization.

Decision Tree Visualization:

Root of tree:



Entire tree: