Imperial College London

COURSEWORK 1: DECISION TREES

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Intro to Machine Learning

Authors: Aidan Holmes, Meg Howard, Hugo Frelin, Igor Sadalski

Date: November 3, 2023

1 Intro

In the following report, the results of a decision tree algorithm implemented to determine in which room measurements were taken based on 7 WIFI signal strengths are discussed. First, the accuracy, recall, precision, and F1 measure of the **unpruned** decision tree using 10-fold cross-validation on both the clean and noisy datasets was recorded. Then the performances of the **pruned** decision tree algorithm were evaluated, again using a nested 10-fold cross-validation to yield the same metrics.

2 Tree Visualisation

A function was created to visualize the decision tree. As an example, see below for the visualized tree trained on the clean dataset.

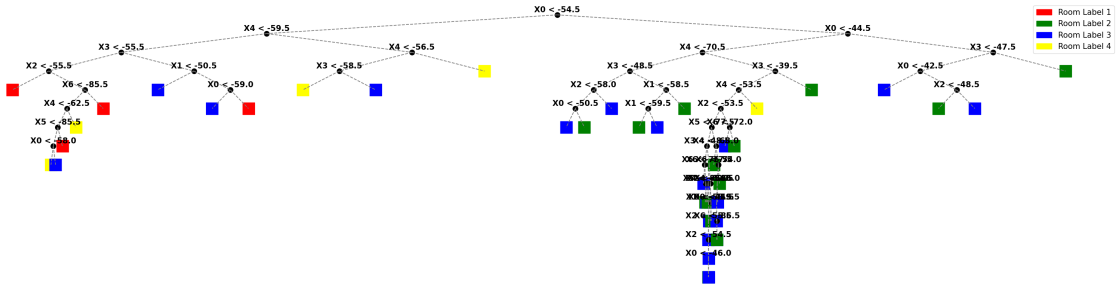

Figure 1: Clean Data Decision Tree

3 Evaluation

3.1 Confusion Matrices

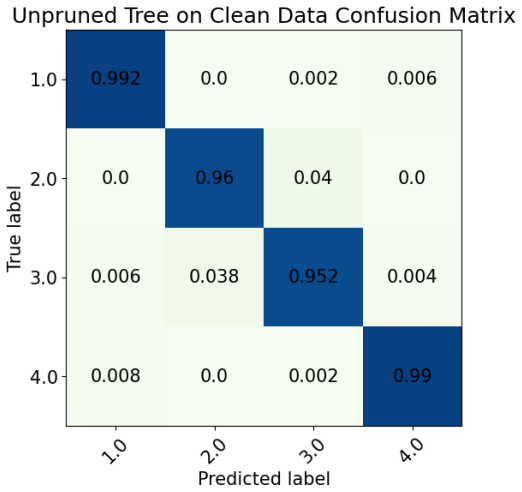

Figure 2: Clean Data Confusion Matrix

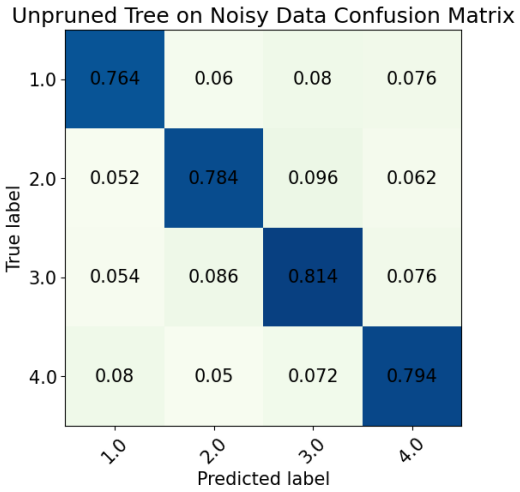

Figure 3: Noisy Data Confusion Matrix

3.2 Performance Metrics For Unpruned Tree

Table 1: Performance Metrics For Unpruned Tree Clean Data

Class	Accuracy	Recall	Precision	F1 Measure
Room 1	99.2%	99.2%	98.6%	98.9%
Room 2	96 %	96 %	96.2%	96.1%
Room 3	95.2%	95.2%	95.6%	95.4%
Room 4	99 %	99 %	99 %	99 %
Macro	97.4%	97.4%	97.4%	97.4%

Table 2: Performance Metrics For Unpruned Tree Noisy Data

Class	Accuracy	Recall	Precision	F1 Measure
Room 1	76.4%	78 %	80.4%	79.2%
Room 2	78.4%	78.9%	80 %	79.4%
Room 3	81.4%	79 %	76.6%	77.8%
Room 4	79.4%	79.7%	78.8%	79.2%
Macro	78.9%	78.9%	79 %	78.9%

3.3 Result analysis

In the noisy data, Room 3 has the highest accuracy but the lowest F1-score, indicating many false positives. Despite this, Room 3 also has the lowest accuracy in the clean data, possibly due to data imbalances. Room 1 is the least accurate in the noisy data, likely because it encroaches into the space of other rooms, making it easier to misclassify as it will share similar values to parts of all rooms.

3.4 Dataset differences

The clean dataset shows consistently higher accuracy, recall, precision, and F1 Measures for all rooms and macro averages compared to the noisy dataset. These inaccuracies will be introduced by the more inconsistent and noisy dataset, impairing the performance of our decision tree classifier. The reduced quality of the noisy data also makes it more difficult for the tree to identify meaningful patterns.

4 Pruning

4.1 Tree Pruning Visualisation

In Figure 4 the decision tree before any pruning operation can be seen. Here it is clearly quite dense with a lot of leaves. Then, in Figure 5, we can see the effects of

the pruning algorithm, as the tree drops all of the overfitted leaves.

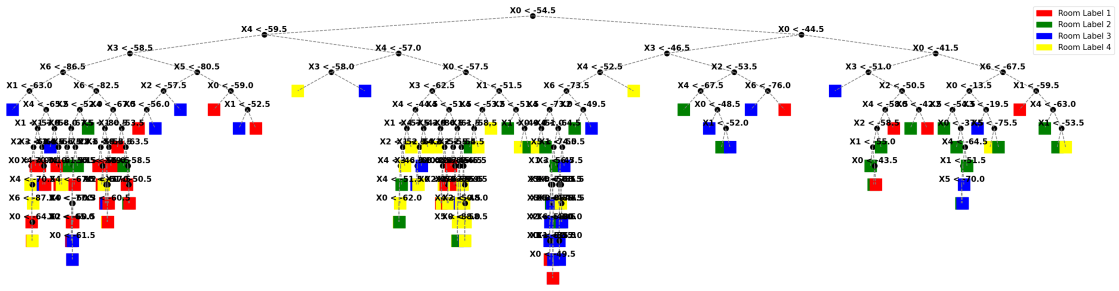

Figure 4: Tree Before Pruning

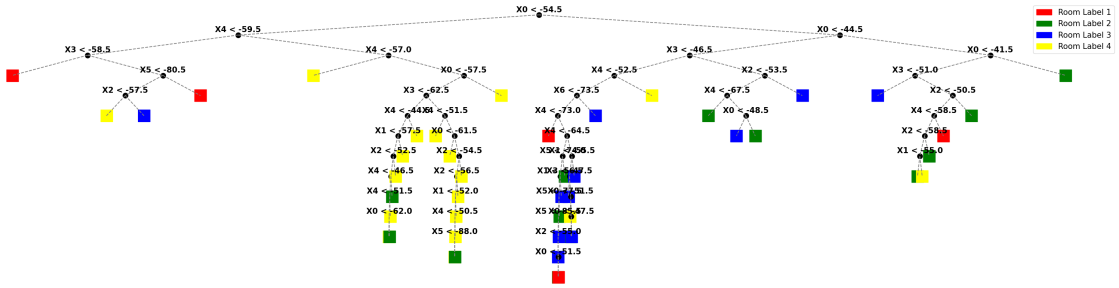

Figure 5: Tree After Pruning

4.2 Confusion Matrix

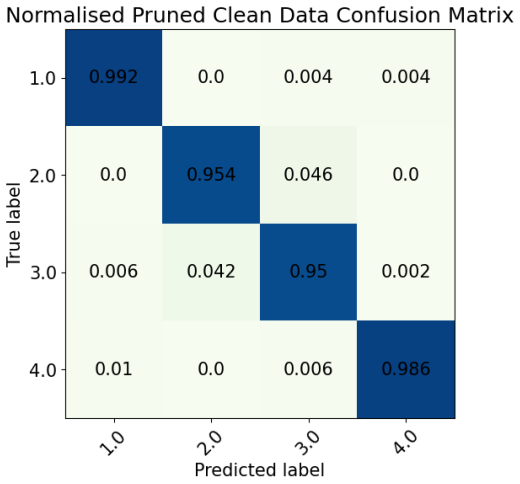

Figure 6: Pruned Clean Data Confusion Matrix

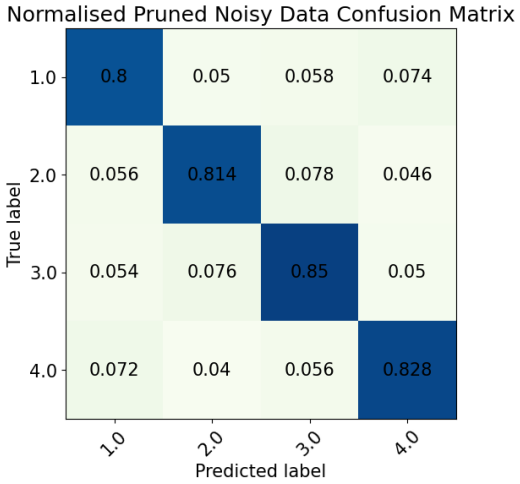

Figure 7: Pruned Noisy Data Confusion Matrix

4.3 Performance Metrics For Pruned Tree

Find below the performance metrics, both by class and the macro values, for the pruned tree on both the clean and noisy datasets.

Table 3: Performance Metrics For Pruned Tree Clean Data

Class	Accuracy	Recall	Precision	F1 Measure
Room 1	99 %	99.2%	98.4%	98.8%
Room 2	95 %	95.4%	95.8%	95.6%
Room 3	95 %	95 %	94.4%	94.7%
Room 4	99 %	98.4%	99.4%	99 %
Macro	97 %	97 %	97 %	97 %

Table 4: Performance Metrics For Pruned Tree Noisy Data

Class	Accuracy	Recall	Precision	F1 Measure
Room 1	80 %	81.5%	81.5%	81.5%
Room 2	81 %	81.9%	83.1%	82.5%
Room 3	85 %	82.5%	81.6%	82 %
Room 4	83 %	83.1%	83 %	83 %
Macro	82.3%	82.3%	82.3%	82.3%

Table 5: Average Depth Comparison

	Noisy	Clean
Max Depth Unpruned	18.98	12.82
Max Depth Pruned:	18.66	12.24
Avg Depth Unpruned	9.73	5.28
Avg Depth Pruned:	9.70	5.24
Median Avg Depth Pruned	9.18	5.38
Median Avg Depth Unpruned	9.67	5.95

4.4 Result analysis

In the clean data, pruning minimally changes metrics; the initial tree captures patterns effectively. In the noisy data, pruning enhances performance, addressing overfitted noise-based branches for a more generalizable model. Relative metrics remain consistent, e.g room 3 still has the highest accuracy and room 1 the lowest. Room 3's increased relative F1 score indicates a decrease in false positives.

4.5 Depth analysis

The average depth of both unpruned/pruned trees differed by a small margin (0.2 to 0.6 units). For the noisy data, pruning boosted accuracy by ~ 3.4%, implying max tree depth isn't directly proportional to accuracy. For the clean data, performance stayed consistent post-pruning. Figures 4/5 show that pruning removes many nodes, but the deepest parts remain unchanged, perhaps concealing the extent of pruning.