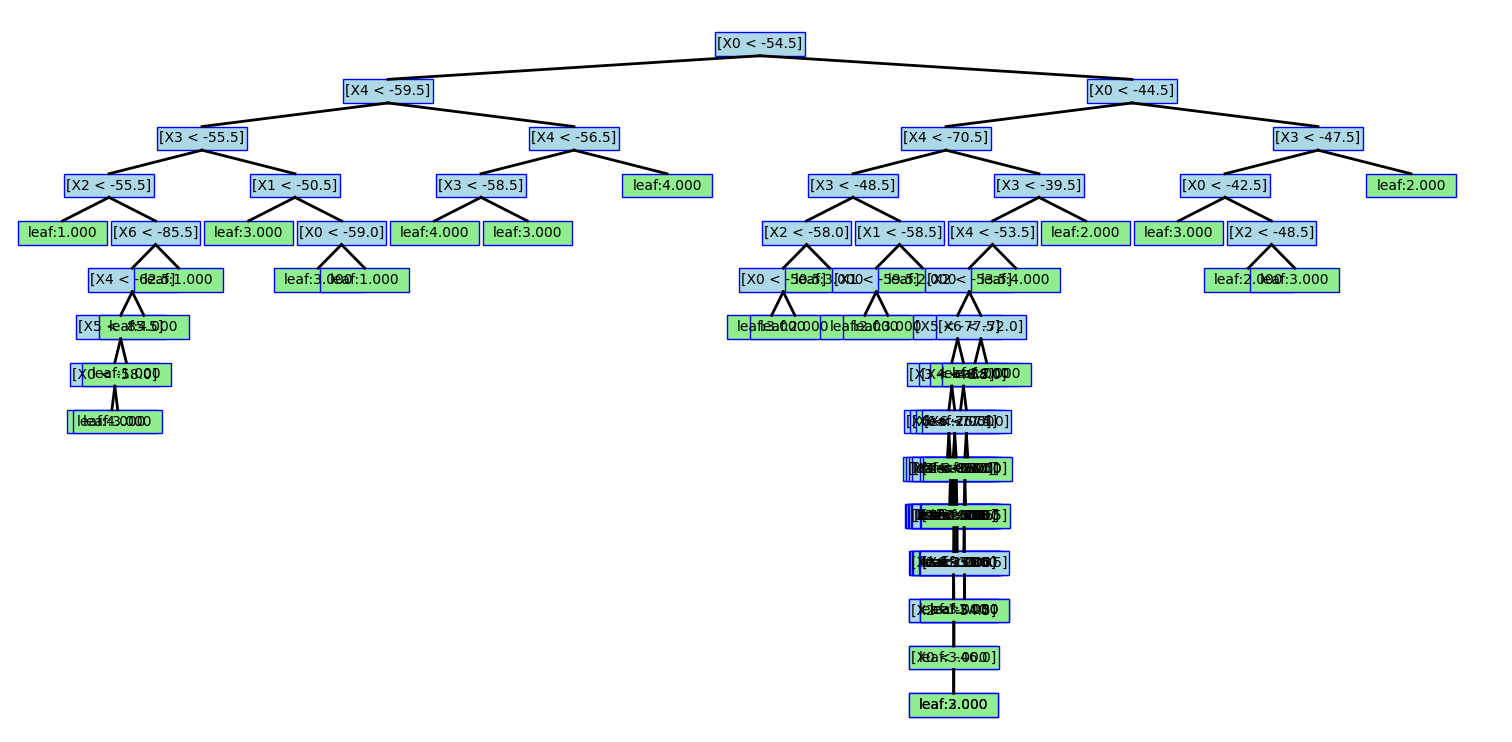
**Step 2: Output of the tree visualisation function**

The visual below shows the decision tree trained on the entire clean dataset. This is the output of from the script decision\_tree.py. And the object method that produces this result is decision\_tree.visualisation()

****

**Step 3: Evaluation**

For this step, we performed evaluation of the decision tree model using a 10-fold cross-validation on both the clean and noisy dataset as required. The evaluation metrics (accuracy, confusion matrix, precision, recall, F1-measures) for the two datasets are presented below. These values came straight from the output of the script evaluation.py. Note that the values in confusion matrix are decimal numbers as they are the average across the 10 folds.

**Cross validation classification metrics:**

**1. Clean Dataset Results:**

Accuracy: 0.972

Confusion Matrix:

[[49.6 0. 0.2 0.2]

[ 0. 47.7 2.3 0. ]

[ 0.1 2. 47.7 0.2]

[ 0.5 0. 0.1 49.4]]

Precision per class: [0.988 0.962 0.95 0.992]

Recall per class: [0.991 0.954 0.956 0.988]

F1-measure per class: [0.989 0.958 0.952 0.99 ]

**2. Noisy Dataset Results:**

Accuracy: 0.807

Confusion Matrix:

[[38.6 3.4 3.7 3.3]

[ 2.7 40.5 4.1 2.4]

[ 2.8 3.4 41.7 3.6]

[ 3. 2.7 3.5 40.6]]

Precision per class: [0.821 0.811 0.787 0.814]

Recall per class: [0.79 0.813 0.811 0.818]

F1-measure per class: [0.803 0.81 0.796 0.815]

**Result analysis:**

In the clean dataset, room1 and room4 are mostly correctly classified. Room2 and room3 are generally classified correctly, with some minor confusion between each other. In the noisy dataset, each room is correctly recognised in the majority of cases. However, each room has significant confusion with all the other rooms. There is considerable confusion between room2 and room3.

**Dataset differences:**

There are significant differences in performance between the two datasets. The accuracy, precision, recall rate, and F1 measure of clean dataset is much higher than that of the noisy dataset. Clean dataset allows for more correct classifications and less confusion between different rooms. The reason is that the noise in the noisy dataset is likely to cause the features of the rooms to overlap, thus more misclassifications between similar rooms.

**Step 4: Pruning and Evaluation**

**Cross validation classification metrics after pruning:**

Note that the values below are rounded to 3 decimal places. They are decimal figures as they are the averaged results.

**1. Clean Dataset Results:**

Accuracy: 0.968

Confusion Matrix:

[[49.644 0. 0.267 0.089]

[ 0. 47.556 2.444 0. ]

[ 0.5 1.956 47.222 0.322]

[ 0.467 0. 0.333 49.2 ]]

Precision per class: [0.981 0.961 0.942 0.992]

Recall per class: [0.992 0.951 0.945 0.982]

F1-measure per class: [0.987 0.955 0.943 0.987]

Average number of nodes before pruning: 73.6

Average tree depth before pruning: 11.656

Average number of nodes after pruning: 28.711

Average tree depth after pruning: 8.167

**2. Noisy Dataset Results:**

Accuracy: 0.879

Confusion Matrix:

[[43.633 1.244 1.756 2.367]

[ 1.889 43.733 2.833 1.244]

[ 2.122 3.233 44.156 1.989]

[ 2.311 1.433 1.811 44.244]]

Precision per class: [0.878 0.883 0.875 0.889]

Recall per class: [0.892 0.881 0.857 0.887]

F1-measure per class: [0.884 0.88 0.865 0.886]

Average number of nodes before pruning: 528.244

Average tree depth before pruning: 18.056

Average number of nodes after pruning: 82.778

Average tree depth after pruning: 13.678

**Result analysis after pruning:**

The accuracy of the clean dataset almost remains unchanged after pruning. For the noisy dataset, pruning improves the prediction accuracy significantly. There are greater numbers of correct classifications for each class and less confusion between them. The differences in the noisy dataset are due to the effect of pruning, which reduces overfitting of our trained model and improves generalisation on the noisy dataset.

**Depth analysis:**

The average depth of the trees for both datasets decreases substantially after pruning, with a reduced number of nodes. As the maximal depth decreases, the prediction accuracy increases. This shows that reducing the tree complexity can lead to improved accuracy, particularly in the noisy dataset, which is more prone to overfitting of the model.