Course Work 1

Decision Trees

Edward Hartley
Jordan Hall
Kert Laansalu
Konstantinos Koupepas

A COMP70050 Course Work Assignment

Imperial College London

November 1, 2022

1 10-Fold Cross-Validation Metrics

1.1 Clean Data

1.1.1 Confusion Matrix

Non-Pruned Decision Tree:

true\pred	1	2	3	4
1	495	0	2	3
2	0	478	22	0
3	3	20	474	3
4	4	0	2	494

Pruned Decision Tree:

true\pred	1	2	3	4
1	497	0	3	0
2	0	470	30	0
3	4	12	481	3
4	5	0	4	491

1.1.2 Average Accuracy

Non-Pruned Decision Tree:

Accuracy = 0.9705

Pruned Decision Tree:

Accuracy = 0.9695

1.1.3 Average Recall Per Class

	Class			
Pruned	1	2	3	4
No	0.99	0.956	0.948	0.988
Yes	0.994	0.94	0.962	0.982

1.1.4 Average Precision Per Class

	Class			
Pruned	1	2	3	4
No	0.98605578	0.95983936	0.948	0.988
Yes	0.98221344	0.97510373	0.92857143	0.99392713

1.1.5 Average F1 Measure Per Class

	Class			
Pruned	1	2	3	4
No	0.98802395	0.95791583	0.948	0.988
Yes	0.98807157	0.95723014	0.94499018	0.98792757

1.2 Noisy Data

1.2.1 Confusion Matrix

Non-Pruned Decision Tree:

true\pred	1	2	3	4
1	395	27	33	35
2	36	382	51	28
3	32	37	413	33
4	45	25	30	398

Pruned Decision Tree:

true\pred	1	2	3	4
1	430	14	19	27
2	18	434	34	11
3	20	50	428	17
4	26	21	24	427

1.2.2 Average Accuracy

Non-Pruned Decision Tree:

Accuracy = 0.794

Pruned Decision Tree:

Accuracy = 0.8595

1.2.3 Average Recall Per Class

	Class			
Pruned	1	2	3	4
No	0.80612245	0.76861167	0.80194175	0.79919679
Yes	0.87755102	0.87323944	0.83106796	0.85742972

1.2.4 Average Precision Per Class

	Class			
Pruned	1	2	3	4
No	0.77755906	0.81104034	0.78368121	0.80566802
Yes	0.87044534	0.83622351	0.84752475	0.88589212

1.2.5 Average F1 Measure Per Class

	Class			
Pruned	1	2	3	4
No	0.79158317	0.7892562	0.79270633	0.80241935
Yes	0.87398374	0.85433071	0.83921569	0.87142857

2 Non-Pruned Results Analysis

In the clean dataset, it is clear that rooms 2 and 3 are often confused with one another frequently. It is also the case that classes 1 and 4 are recognized with very high accuracy according to the F1 measure. In the noisy dataset, all classes are recognized with lower accuracy and smaller differences between them, however, classes 1 and 4 still have higher F1 measures than classes 2 and 3, which are still confused at a slightly higher rate.

3 Dataset differences

We can clearly see that the average accuracy of the clean and noisy datasets are very different, going from 97.05% to 79.4%. This is mainly because Decision Trees will perfectly fit to training datasets, which causes them to overfit and do poorly on test data.

4 Pruned Results Analysis

On the clean dataset, pruned Decision Trees do worse than the non-pruned Decision Tree, with accuracy 97.05% and 96.95% respectively. This is because the validation set we're pruning with may lack samples that the non-pruned Decision Tree fit perfectly too. This causes a decrease in test performance as a result.
On the noisy dataset, the pruned Decision Tree outperforms the non-pruned Decision Tree, with accuracy of 79.4% and 85.95% respectively. This is because the pruned Decision Trees help mitigate overfitting on the noisy dataset.

5 Depth Analysis

Average Depths of Pruned/Non-Pruned Decision Trees By Dataset:

	Clean	**Noisy**
Non-Pruned	13.5	18.8
Pruned	7.8	10.9

As we can see by the above table and the cross-validation metrics report, larger trees do better on clean datasets, and worse on noisy datasets. This makes sense when you consider that clean datasets do not have outliers, and overfitting is not a problem. However, on noisy datasets, overfitting (having a larger tree) can overfit for outlier data points.