1. Performance metrix:

Dataset	Accuracy	Precision(weighted)	Recall(weighted)	F1(weighted)
Mushrooms.csv	1.0	1.0	1.0	1.0
Nursery.csv	0.98	0.98	0.98	0.98
Tictactoe.csv	0.88	0.88	0.88	0.88

The **mushrooms** dataset achieved 100% across all metrics, showing that it is well-fit and clean dataset. The model has made all correct predictions which meant that It has found all poisonous mushrooms

The **Nursery** dataset also performed quite well, with close to around 99% accuracy and this tells us that the dataset is almost perfectly-fit and clean and has a little bit of noise present in it. Most of the predictions made were correct but it has made quite a few miss-predictions.

The **tic-tac-toe** Lower performance compared to other. This tells us that dataset isn't clean as the other two and that its accuracy can be improved.

2. Tree Characteristics Analysis

Dataset	Max dept	Total node	Leaf node	Internal node
Mushrooms	4	29	24	5
Nursery	7	983	703	280
Tic-tac-toe	7	260	165	95

Mushroom.csv: simplest and most shallow

Nursery.csv: it is the most largest due to multiple target classes

Tictactoe.csv- it is quite large, make accurate predictions if we capture complex patterns

3. Dataset-Specific Insights

Mushroom.csv: Strong single attribute almost perfect separate classes

Nursery.csv: Many attributes interact

Tictactoe- Features are symmetric, the root node often focuses on the center pos

4. Comparative Analysis

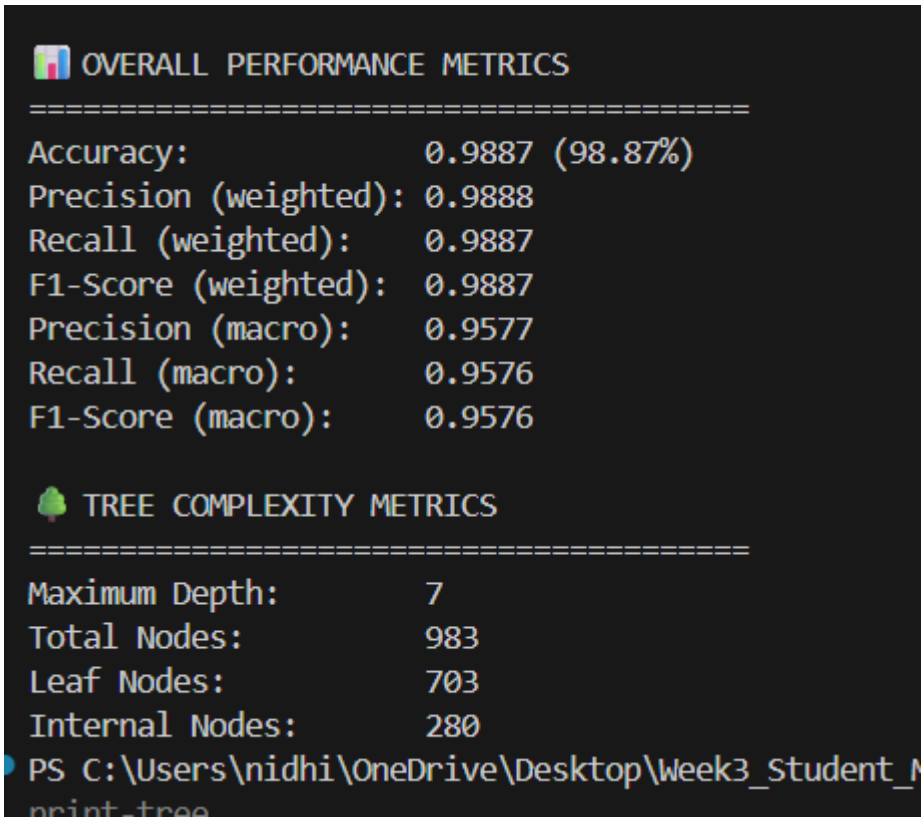
Alg performance:

a. Mushroom dataset had the highest accuracy

b. Largest dataset increases the number of test cases to be run and hence the accuracy improves, smallest dataset may result in lower accuracy
c. More categorical levels produced larger trees.
d. Nursery is affected by imbalance, mushroom is fairly balanced and tic-tac-toe is binary so less affected.

5. Improvement:
    a. Apply pruning to reduce overfitting and simplify complex tree
    b. Mushroom: no change is needed
    c. Nursery: fix the class imbalance
    d. Tic-tac-toe: use a diff alg

Nursery.csv



tictactoe

```
📊 OVERALL PERFORMANCE METRICS
========================================
Accuracy:              0.8836 (88.36%)
Precision (weighted):  0.8827
Recall (weighted):     0.8836
F1-Score (weighted):   0.8822
Precision (macro):     0.8784
Recall (macro):        0.8600
F1-Score (macro):      0.8680

🌳 TREE COMPLEXITY METRICS
========================================
Maximum Depth:         7
Total Nodes:           260
Leaf Nodes:            165
Internal Nodes:        95
```

Mushrooms.csv

```
📊 OVERALL PERFORMANCE METRICS
========================================
Accuracy:              1.0000 (100.00%)
Precision (weighted):  1.0000
Recall (weighted):     1.0000
F1-Score (weighted):   1.0000
Precision (macro):     1.0000
Recall (macro):        1.0000
F1-Score (macro):      1.0000

🌳 TREE COMPLEXITY METRICS
========================================
Maximum Depth:         4
Total Nodes:           29
Leaf Nodes:            24
Internal Nodes:        5
```