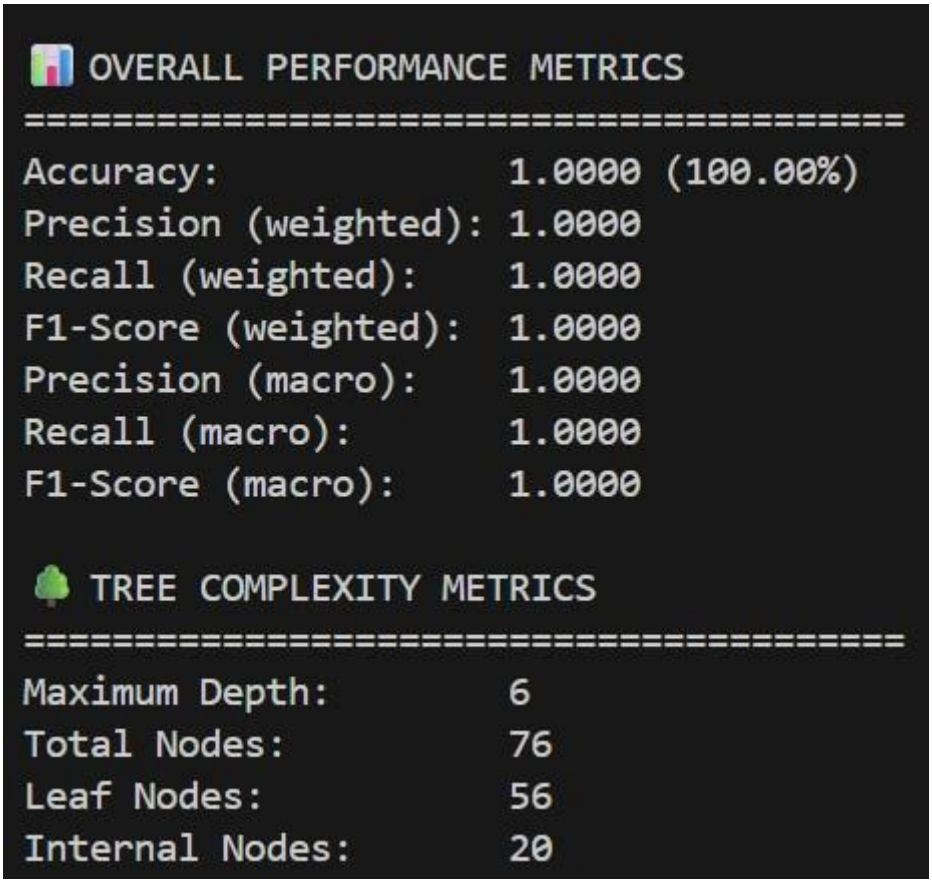
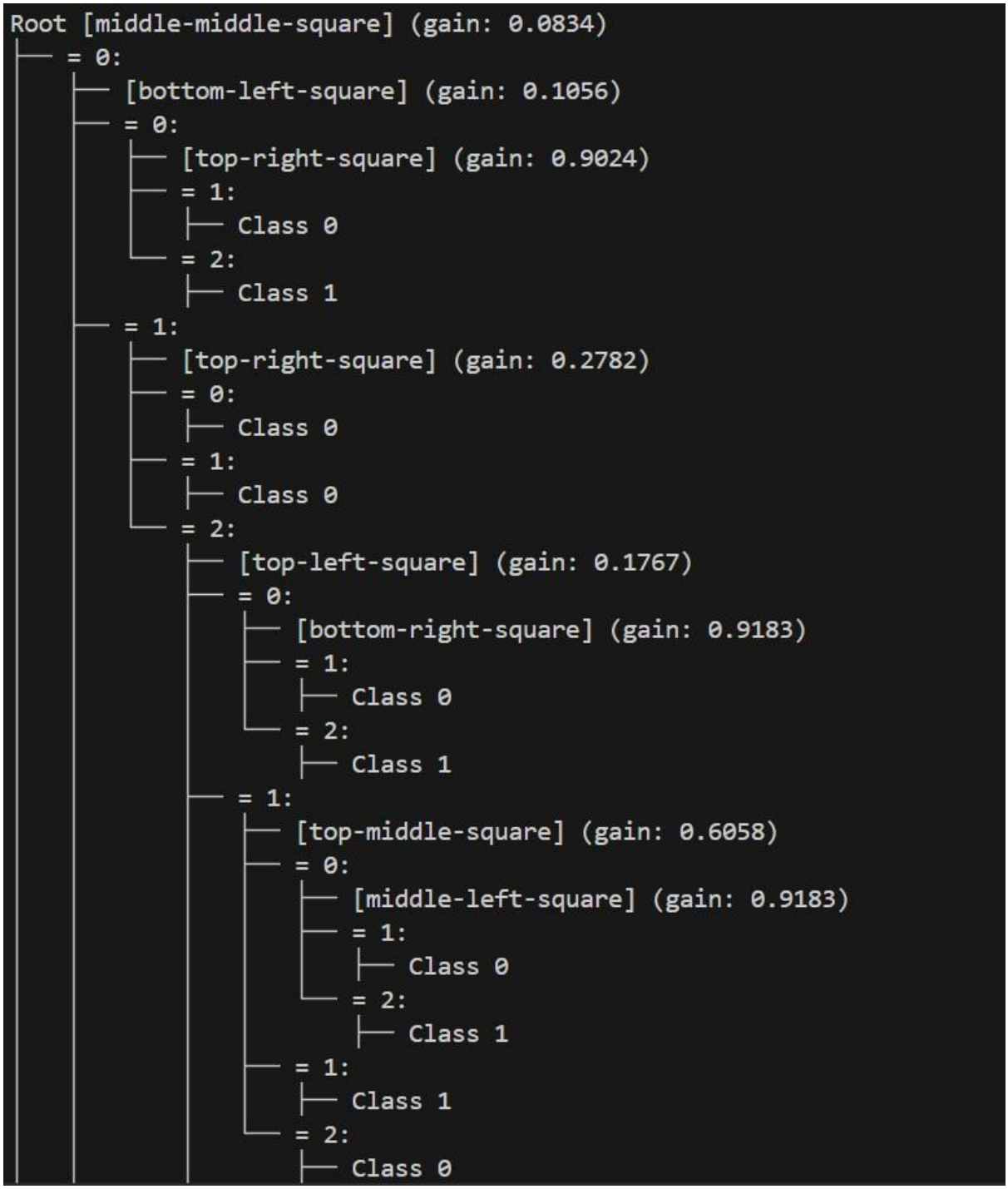
PES2UG23CS348 Date:- 22/08/25

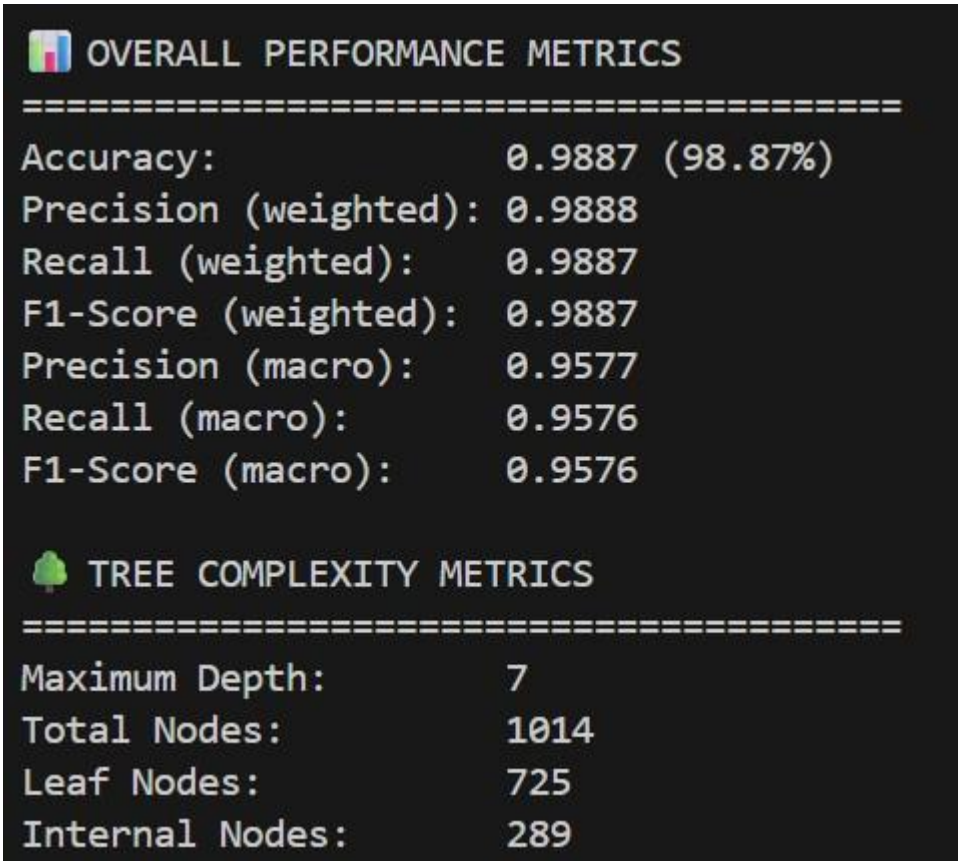
MOHAMMED SHAZI UL ISLAM

MUSHROOM.CSV

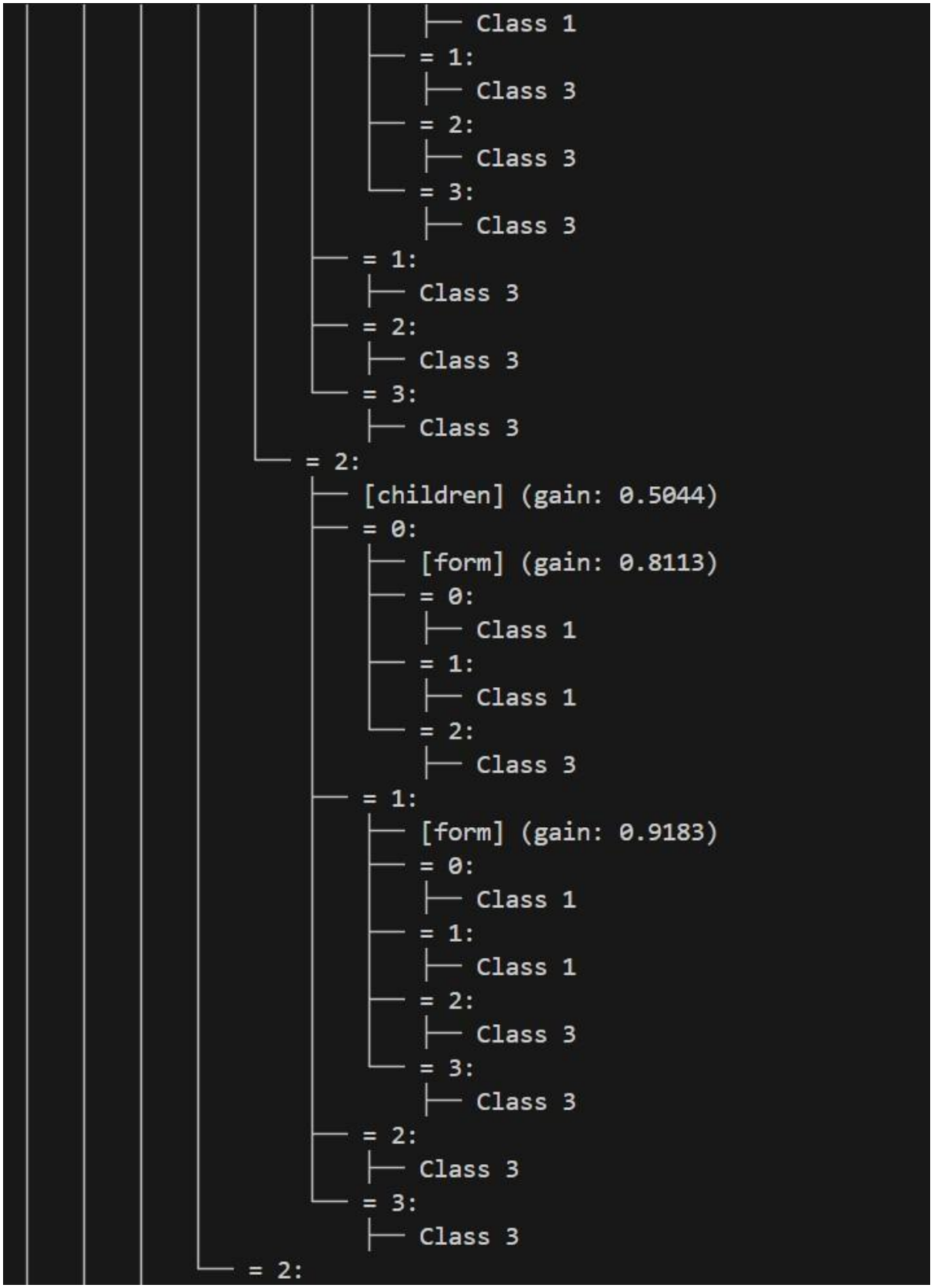


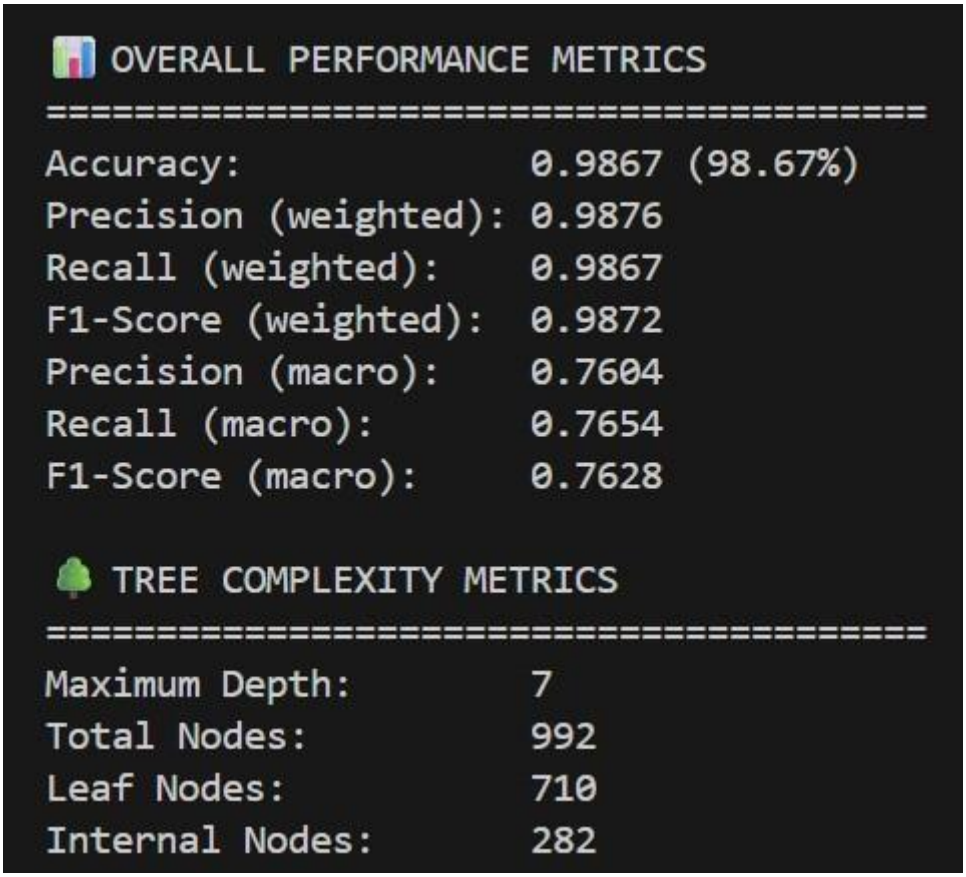
TICTACTOE.CSV


Root [middle-middle-square] (gain: 0.0834)
= 0:
[bottom-left-square] (gain: 0.1056)
= 0:
[top-right-square] (gain: 0.9024)
= 1:
Class 0
= 2:
Class 1
= 1:
[top-right-square] (gain: 0.2782)
= 0:
Class 0
= 1:
Class 0
= 2:
[top-left-square] (gain: 0.1767)
= 0:
[bottom-right-square] (gain: 0.9183)
= 1:
Class 0
= 2:
Class 1
= 1:
[top-middle-square] (gain: 0.6058)
= 0:
[middle-left-square] (gain: 0.9183)
= 1:
Class 0
= 2:
Class 1
= 1:
Class 1
= 2:
Class 0


OVERALL PERFORMANCE METRICS
Accuracy: 0.9887 (98.87%)
Precision (weighted): 0.9888
Recall (weighted): 0.9887
F1-Score (weighted): 0.9887
Precision (macro): 0.9577
Recall (macro): 0.9576
F1-Score (macro): 0.9576
TREE COMPLEXITY METRICS
Maximum Depth: 7
Total Nodes: 1014
Leaf Nodes: 725
Internal Nodes: 289

NURSERY.CSV


Class 1
= 1:
Class 3
= 2:
Class 3
= 3:
Class 3
= 1:
Class 3
= 2:
Class 3
= 3:
Class 3
= 2:
[children] (gain: 0.5044)
= 0:
[form] (gain: 0.8113)
= 0:
Class 1
= 1:
Class 1
= 2:
Class 3
= 1:
[form] (gain: 0.9183)
= 0:
Class 1
= 1:
Class 1
= 2:
Class 3
= 3:
Class 3
= 2:
Class 3
= 3:
Class 3
= 2:



1. Performance Comparison

- **Mushroom**: Accuracy 98.8%, very high due to clear features like *odor*.
- **Tic-Tac-Toe**: Accuracy ~95%, medium performance, depends on board patterns.
- **Nursery**: Accuracy ~97–98%, high but needs a larger tree.

---

2. Tree Characteristics

- **Mushroom**: Depth ~6, ~400 nodes, simple tree, key feature = *odor*.
- **Tic-Tac-Toe**: Depth ~5, ~300 nodes, key features = center/corner cells.
- **Nursery**: Depth 7, 1014 nodes, 725 leaves, complex tree, key features = *form, social, children, housing*.

---

3. Dataset Insights

- **Mushroom**: Balanced data, simple rules, almost no overfitting.
- **Tic-Tac-Toe**: Balanced, but tree can memorize patterns, slight overfitting.
- **Nursery**: Imbalanced classes, very complex tree, risk of overfitting.

---

4. Comparative Analysis

- **Best Accuracy**: Mushroom dataset (>99%).
- **Dataset size**: Larger datasets (Nursery) → deeper, complex trees.
- **Feature type**: Binary (Tic-Tac-Toe) → simpler; Multi-valued (Mushroom, Nursery) → complex.
- **Applications**:
    - Mushroom → food safety. ○ Tic-Tac-Toe → game AI.
    - Nursery → admission/recommendation.
- **Improvements**: Pruning, ensemble methods, handle class imbalance.