REPORT WEEK-3 ML LAB

Comparative Analysis of ID3 Decision Tree on 3 datasets-mushrooms, tic-tac-3oe, and Nursery

Analysis:

Performance Comparison

MUSHROOM: the PyTorch implementations gave 100% accuracy. Precision, recall, and F1-score were also perfect because the attribute “odor” almost completely determines the class.

TIC-TAC-TOE: PyTorch achieved 87.30% accuracy. Precision, recall, and F1-scores were close to each other but lower compared to mushroom, since the dataset is harder and involves many different board positions.

NURSERY: PyTorch reached 98.67% accuracy. Even though the dataset is large and has many classes, the tree still classified very well. Precision and recall were strong for the common classes but weaker for the rare ones.

Tree Characteristics Analysis

MUSHROOM: tree was shallow, usually 3–5 levels, with few nodes. The key attribute was odor, which almost completely split the data.

TIC-TAC-TOE: trees were deeper, about 7–9 levels, since all nine board cells matter. The middle position was often selected first because of its importance in deciding the outcome

NURSERY: deepest and largest trees, usually above 10 levels. Features like parents and finance were chosen early, and the tree size reflected the dataset’s complexity and many attribute values.

Dataset-Specific Insights

MUSHROOM: Odor was the most important feature. The dataset was balanced, and rules were simple such as “if odor = foul then poisonous.” Overfitting was not a problem.

TIC-TAC-TOE: Central and diagonal cells were most important. The dataset was balanced, but the tree sometimes overfit by memorizing board states instead of generalizing.

NURSERY: Parents and finance features were dominant. The dataset was imbalanced with “not_recom” as the majority class. There were signs of overfitting.

Comparative Analysis

a) Algorithm Performance:

a. Which dataset achieved the highest accuracy and why?

Mushroom (100%) dataset achieved the highest accuracy. The key reason is a single, highly discriminative attribute—odor—that almost perfectly separates edible from poisonous classes, yielding perfect precision/recall/F1 as well.

Mushroom (~100%) > Nursery (~98.7–98.9%) > Tic-Tac-Toe (~87–88%).

b. How does dataset size affect performance?

Larger datasets like Nursery in this case, improve generalization but produce deeper, more complex trees. Despite many classes and high depth, accuracy remains high because abundant examples help ID3 learn robust splits.

Smaller datasets like the Tic-Tac-Toe here, cap peak accuracy; with many distinct board states relative to sample count, trees may overfit specific patterns instead of general strategies

c. What role does the number of features play?

When we have fewer but "strong" features (Mushroom; e.g., odor) ,this can lead to shallow trees (≈3–5 levels) and excellent accuracy.

And in case of many features or many-valued categorical features (Nursery) increase tree depth (often >10) and node count, raising complexity while still performing well due to data volume.

Lastly in case of binary features across many positions (Tic-Tac-Toe's nine cells) create moderately deep trees (≈7–9 levels) and more complex decision paths than Mushroom

b)Data Characteristics Impact:

• How does class imbalance affect tree construction?

Among the given the datasets, Nursery is imbalanced (majority class: not_recom). ID3 focuses early splits on features that best separate the dominant class, which can inflate overall accuracy yet depress precision/recall for minority classes; deeper trees show signs of overfitting on rare classes.

•Which types of features (binary vs multi-valued) work better?

Binary features which are available in the dataset tic-tac-toe, provide clear, simple splits, but with many positions/states can still yield sizable depth and moderate accuracy.

Multi-valued categorical features that are available in the Nursery dataset tend to increase branching and depth, complicating interpretability even when accuracy is high.

By contrast, Mushroom's categorical features include one highly informative attribute, making it both accurate and simple.

c)Practical Applications:

- **For which real-world scenarios is each dataset type most relevant?**
    - Mushroom: Systems needing high-stakes, high-interpretability decisions, e.g., food safety/toxicology identification or other domains where one or two features dominate outcomes.
    - Tic-Tac-Toe: Pedagogical/game AI demonstrations, strategy learning on grid-like states, or as a sandbox for evaluating overfitting and pruning techniques.
    - Nursery**:** Decision support systems (admissions/recommendations/eligibility) with many stakeholder criteria and class labels.
- **What are the interpretability advantages for each domain?**
    - Mushroom: Best interpretability—shallow tree; rules like "if odor=foul ,then it is poisonous" are transparent and actionable.
    - Tic-Tac-Toe: Moderate interpretability—you can trace plays (center/diagonals appear early), but paths are longer.
    - Nursery: Lowest interpretability—deep, large trees with many-valued splits; understanding global logic is harder.
- **How would you improve performance for each dataset?**
    - Mushroom: Already perfect—no model changes needed.
    - Tic-Tac-Toe : Mitigate overfitting and boost generalization with pre-pruning/post pruning or feature engineering to capture tactical patterns.
    - Nursey: address imbalance and complexity with resampling/class-weighted splits or pruning to reduce depth and improve interpretability.