Decision Trees

**Tree Depth**

Let me explain why choosing a max\_depth of 3 is often a good choice for decision trees, especially with this weather dataset:

1. \*\*Interpretability\*\*:

- A depth of 3 means at most 3 decision nodes from root to any leaf

- This creates a tree with at most 8 leaf nodes (2³)

- Humans can easily follow and understand trees of this depth

- It's like asking at most 3 questions before making a decision (e.g., "Is it sunny?" → "Is it hot?" → "Is it humid?")

2. \*\*Preventing Overfitting\*\*:

- Deeper trees can memorize the training data too well

- With your weather dataset (14 samples), a deeper tree could create a path for each specific case

- This would give 100% accuracy on training data but might perform poorly on new data

- max\_depth=3 forces the tree to find general patterns instead of memorizing

3. \*\*Data Size\*\*:

- Your dataset has 14 examples

- With 4 features (Outlook, Temperature, Humidity, Wind)

- A deeper tree would have too few samples at the leaf nodes

- Rule of thumb: each leaf should have enough samples to make reliable predictions (typically 5+ samples)

4. \*\*Feature Count\*\*:

- You have 4 features

- max\_depth=3 means you can use up to 3 of these features

- This is reasonable since not all features may be equally important

- The tree will select the most informative features first

Would you like me to demonstrate what happens if we try different max\_depth values with your dataset to see how it affects:

1. Model accuracy

2. Tree complexity

3. Risk of overfitting

4. Interpretability