The ID3 (Iterative Dichotomiser 3) algorithm builds a decision tree top-down by selecting attributes that maximize information gain, using entropy as the impurity measure. Entropy quantifies uncertainty:

$$Entropy\ (S) = -\sum p_i * \log_2(p_i)$$

where p_i is the proportion of each class in set S.

Information gain is Entropy(parent) - weighted average Entropy(children). The process is recursive until subsets are pure (entropy=0) or no attributes remain.

Here are the steps applied to the dataset:

1. Calculate initial entropy: for 10 instances (6 Yes, 4 No)

   Entropy = $-\left(\frac{6}{10}\right)\log_2\left(\frac{6}{10}\right) - \left(\frac{4}{10}\right)\log_2\left(\frac{4}{10}\right) \approx 0.971\ bits$

2. Compute information gain for each attribute

   Outlook ≈ 0.322

   Temperature ≈ 0.095

   Humidity ≈ 0.125

   Wind ≈ 0.091

3. Select root node: Outlook (max gain, 0.322).

   Split dataset into subsets:

   Sunny – 4 instances, entropy 0.811

   Overcast – 2 instances, entropy 0

   Rain – 4 instances, entropy 0.811

4. Handle pure subsets: Overcast is pure (all Yes), leaf node: Yes
5. Recurse on impure subsets

   For Sunny: Gains – Temperature ≈ 0.811, Humidity ≈ 0.811, Wind ≈ 0.123. Select Humidity (ties broken by simplicity/fewer branches).

   - Subsets: High (3 No, entropy=0 → No), Normal (1 Yes, entropy=0 → Yes).

   For Rain: Gains - Temperature ≈ 0.311, Humidity ≈ 0.123, Wind ≈ 0.811 (highest). Select Wind.

- Subsets: Weak (3 Yes, entropy=0 → Yes), Strong (1 No, entropy=0 → No).

6. Stop recursion: All leaves are pure. No need of pruning anymore.

Decision Tree

The tree predicts ‘play cricket’ based on weather.