**Full Decision Tree Construction Algorithm (ID3):**

1. Each node \( v \): has a subset of training set \( S(v) \)
2. Initially \( v \) root; \( S(\text{root}) = \) full training set.

1. While \( \exists \) an "impure" leaf of the current tree:
   - Find a splitting rule for \( v \). Say rule is: "Is \( x_f \leq t \)?"
   - Replace \( v \) with:

   ![Diagram]

   
   \[ S(u_1) = \text{Subset of } S(u) \text{ for which } x_f \leq t. \]
   
   \[ S(u_2) = S(u) \setminus S(u_1) \]

   - If either \( u_1 \) or \( u_2 \) are "pure", make them a prediction node (that predicts the label of all of \( S(u_1) \) or \( S(u_2) \)).

* "Pure" node: all data points associated with this node have the same label. Impure: Not pure.

---

**Splitting Rule:**

1. For all pairs of features \( f \), thresholds \( t \):
   
   * Compute Information Gain
   
   \[ IG = H(X) - H(Z|Z) \]

   * Pick the pair that maximizes the Information Gain.