1. **What are impurity measures in Decision Trees?**
Impurity measures quantify the disorder in data at a node. Common measures include Gini Impurity, which evaluates the probability of incorrect classification, and Entropy, which measures information randomness. Lower impurity indicates better splits, leading to purer child nodes and an improved Decision Tree.

2. **What is the mathematical formula for Gini Impurity?**
Gini Impurity is calculated as: Gini=1−i=1∑npi2
where pip_i is the probability of class ii. A lower Gini value indicates a purer node.

3. **What is the mathematical formula for Entropy?**
Entropy is defined as:
Entropy=−i=1∑npilog2pi
where pip_i is the probability of class ii. Higher entropy means more uncertainty.

4. **What is Information Gain, and how is it used in Decision Trees?**
Information Gain measures the reduction in entropy after a split. It is calculated as:
IG=Entropyparent−i∑(NNi×Entropychild)Higher Information Gain indicates a better split, guiding tree growth.

5. **What is the difference between Gini Impurity and Entropy?**
Gini Impurity focuses on misclassification probability, while Entropy measures information randomness. Gini is computationally simpler, whereas Entropy is more sensitive to changes in class distribution. Both are used for selecting the best split in Decision Trees.

6. **What is the mathematical explanation behind Decision Trees?**
Decision Trees use recursive binary splitting. At each node, a feature is chosen based on a criterion (Gini/Entropy) that minimizes impurity. The process continues until stopping conditions are met, forming a tree with decision rules at internal nodes and final classifications at leaves.

7. **What is Pre-Pruning in Decision Trees?**
Pre-Pruning stops tree growth early based on constraints like maximum depth, minimum samples per split, or minimum impurity decrease. It prevents overfitting by restricting complexity but risks underfitting if applied too aggressively.

8. **What is Post-Pruning in Decision Trees?**
Post-Pruning removes branches from a fully grown tree using validation data. It reduces

overfitting by eliminating less significant splits, improving generalization while preserving essential patterns in the data.

9. **What is the difference between Pre-Pruning and Post-Pruning?**
Pre-Pruning stops tree growth early based on predefined constraints, whereas Post Pruning removes unnecessary branches after full growth. Pre-Pruning prevents overfitting but may underfit, while Post-Pruning improves generalization without premature stopping.

10. **What is a Decision Tree Regressor?**
A Decision Tree Regressor predicts continuous values instead of class labels. It partitions data into smaller regions based on feature values and assigns the mean or median of training points in each region as the prediction.

11. **What are the advantages and disadvantages of Decision Trees?**
Advantages: Simple to understand, interpretable, handles numerical and categorical data, and requires minimal preprocessing.
Disadvantages: Prone to overfitting, sensitive to small variations in data, and biased towards dominant classes.

12. **How does a Decision Tree handle missing values?**
Decision Trees handle missing values using surrogate splits, ignoring missing data during split calculations, or replacing missing values with the most common/frequent value of the feature.

13. **How does a Decision Tree handle categorical features?**
Categorical features are handled by converting them into numerical values using one-hot encoding or label encoding. The tree then selects splits based on impurity reduction, treating categorical values as separate branches.

14. **What are some real-world applications of Decision Trees?**
Decision Trees are used in medical diagnosis, fraud detection, customer segmentation, recommendation systems, and credit risk analysis. They provide clear decision-making rules, making them valuable for interpretability in business and healthcare applications.