1.What is the estimated depth of a Decision Tree trained (unrestricted) on a one million instance training set?

Answer: The estimated depth of a Decision Tree trained on a one million instance training set depends on the complexity of the problem being solved and the level of noise in the data. In general, a Decision Tree trained on a large training set is likely to have a larger depth to capture the complex relationships between the features and the target variable.

2.Is the Gini impurity of a node usually lower or higher than that of its parent? Is it always lower/greater, or is it usually lower/greater?

Answer: The Gini impurity of a node is usually lower than that of its parent, as splitting a node based on a feature reduces the impurity of the resulting child nodes. However, it is not always the case, as splitting a node on a noisy or irrelevant feature may increase the impurity of the resulting child nodes.

3.Explain if it's a good idea to reduce max depth if a Decision Tree is overfitting the training set?

Answer: Yes, it's a good idea to reduce the max depth of a Decision Tree if it's overfitting the training set. Overfitting occurs when a Decision Tree is too complex and captures the noise in the training data. By reducing the max depth, the Decision Tree becomes simpler and less prone to overfitting, leading to better generalization performance on new data.

4.Explain if it's a good idea to try scaling the input features if a Decision Tree underfits the training set?

Answer: No, it's not necessary to scale the input features of a Decision Tree if it's underfitting the training set. Decision Trees are insensitive to the scale of the input features and can handle them without the need for scaling. Underfitting occurs when a Decision Tree is too simple and fails to capture the underlying patterns in the training data. In this case, it's better to increase the complexity of the Decision Tree by increasing the max depth or adding more features.

5.How much time will it take to train another Decision Tree on a training set of 10 million instances if it takes an hour to train a Decision Tree on a training set with 1 million instances?

Answer: The time required to train a Decision Tree on a training set depends on various factors such as the complexity of the problem, the number of features, and the hyperparameters used. In general, training a Decision Tree on a training set with 10 million instances may take approximately 10 hours if it takes an hour to train a Decision Tree on a training set with 1 million instances.

6.Will setting presort=True speed up training if your training set has 100,000 instances?

Answer: No, setting presort=True is not likely to speed up training if your training set has 100,000 instances. Presorting the data can be beneficial for small datasets but can become computationally expensive for larger datasets. In general, it's recommended to leave presort=False for larger datasets.