

# Top 10 Must Read Interview Questions on Decision Trees

[ALGORITHM](#)[BEGINNER](#)[CLASSIFICATION](#)[INTERVIEW QUESTIONS](#)[INTERVIEWS](#)[MACHINE LEARNING](#)

## Introduction

Decision trees are one of the most widely used algorithms in machine learning which provide accurate and reliable results that can be used for classification and regression problems. In data science interviews, questions are mostly asked related to decision trees. It is essential to answer those questions efficiently as it is one of the most useful, necessary, and basic algorithms.

This article will discuss the [top questions](#) related in [machine learning interviews](#) and their appropriate solutions. This will help one to answer those questions efficiently and appropriately.

### Learning Objectives

After going through this article, you will learn:

1. The terminologies CART and ID3 algorithms.
2. Understanding of entropy and information gain.
3. The properties of decision trees.
4. Overfitting and underfitting.
5. A proper way to answer interview questions related to the decision tree.
6. To design and deliver an appropriate answer to interview questions efficiently.

This article was published as a part of the [Data Science Blogathon](#).

## Table of Content

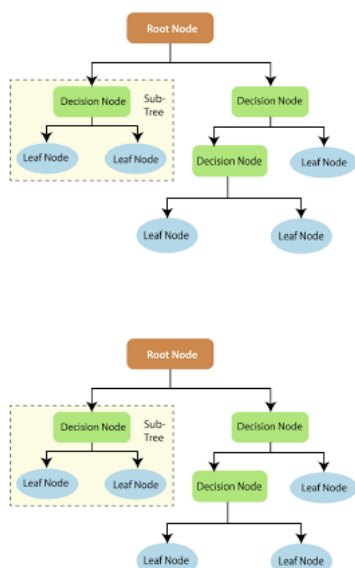
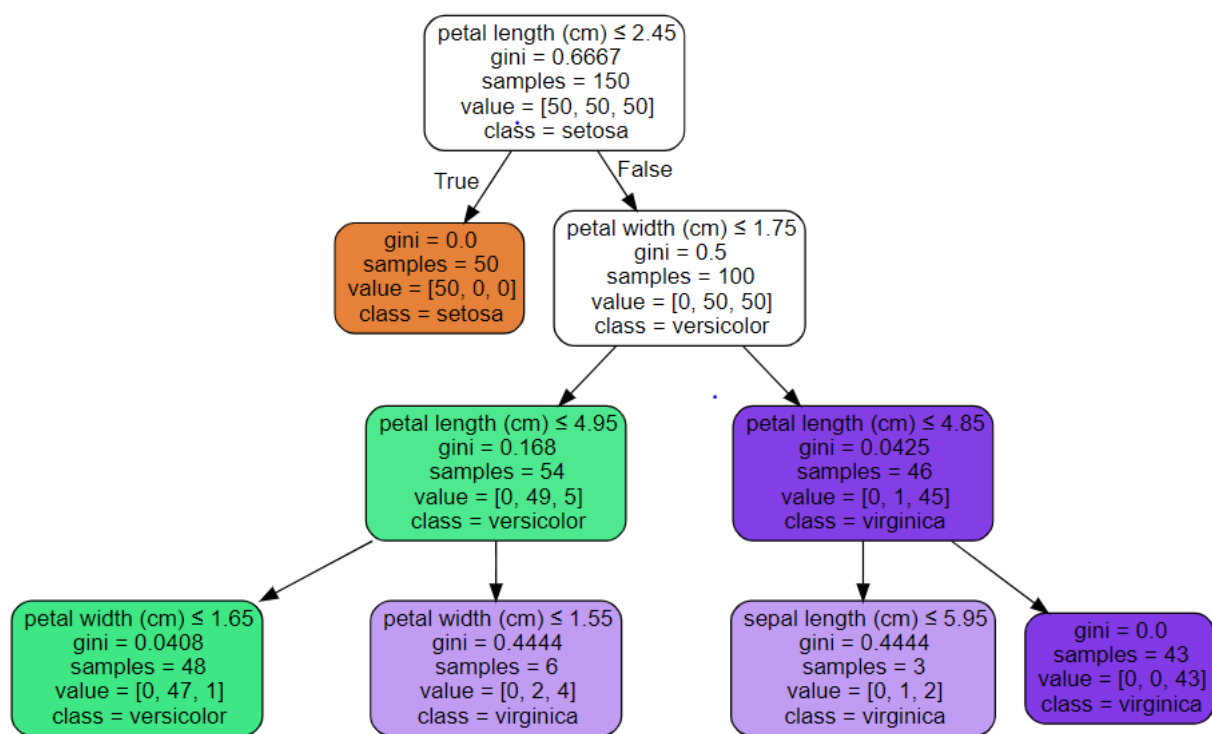
1. What are CART algorithms in machine learning?
2. What are ID3 algorithms in decision trees?
3. What is entropy in decision trees?
4. What is Information gained in decision trees?
5. What is gini entropy in decision trees?
6. What are root and leaf nodes in decision trees?
7. What are shallow and deep decision trees?
8. How do underfitting and overfitting occur?
9. Why are decision trees unstable?
10. Why is feature scaling not required in decision trees?
11. Conclusion

## Q1. What are CART Algorithms in Machine Learning?

[CART](#) stands for Classification and Regression Trees, an algorithm that splits the data from the top to the bottom level based on information gain and entropy. This algorithm calculates the impurity for each possible split and then decides according to the minimum loss.

## Q2. What are ID3 Algorithms in Decision Trees?

ID3 algorithms are a type of algorithm that works on the principle of impurity, and information gain of the split, the algorithm calculates the entropy and information gain for all splits and divides the data accordingly.



Here note that the CART algorithm is used for binary trees, for example, yes or no. At the same time, the ID3 algorithm is used for multiple class nodes where the tree has more than two leaf nodes.

## Q3. What is Entropy in Decision Trees?

Entropy is a measure of uncertainty or impurity that a particular split or node has. Here the data is divided into many divisions, where it can have some impurity. If the data is wholly split homogeneously, the entropy will be zero.

$$\text{Entropy} = -y_i(\log(y_i))$$

## Q4. What is Information Gain?

As the name suggests, the information gained in any decision tree node or split is the information that a particular split provides to the final model. According to the knowledge gained from the split, the last split is a decision on a specific node. The more information gained, the more the chance of implementing the particular split. **Information Gain = E(Parent Node) – Weighted Entropy(Child Node)**

## Q5. What is Gini Entropy?

Gini entropy is also a type of entropy used while splitting the dataset and its nodes. Here the gini entropy is slightly different from the classical entropy, and the calculation of the same is less complex than classical entropy. **Gini Entropy =  $1 - ((P_1)^2 + (P_2)^2)$**

As we can see in the formula calculating the gain entropy for any split is very easy as we need to calculate the squares of the class and then substitute it, whereas in classical entropy, calculating the log of the classes makes the calculation more complex. Hence, the gini entropy is preferred most for the decision tree algorithm.

## Q6. What are Root and Leaf Nodes in Decision Trees?

Root nodes in decision trees are the nodes or points from where the tree starts or where the dataset splits. It is the topmost level of the tree.

Whereas the leaf nodes are that nodes are subparts of splits; in simple words, if the root node is split into two parts, then that particular two parts are known as the leaf nodes of the particular root node.

## Q7. What are Shallow and Deep Decision Trees?

As the name suggests, shallow decision trees are the type of decision trees that are not grown for a very high level of depth and has significantly fewer splits from the data. For example, the decision tree has only three splits. The deep decision trees are grown to their maximum depth and have many splits in the dataset. For example, the decision tree grows to max depth having 15 or 25 splits.

## Q8. How do Overfitting and Underfitting Occur?

In the case of very shallow trees, the tree will have only 2 or 3 splits in the data, meaning that the algorithms will not learn something valuable from the data as there is no sufficient splish in the data. The algorithm will result in poor performing model in training and testing both data, which finally results in the underfitting of the model.

Whereas if we consider a tree grown to its max depth, the tree will know every pattern of the data, and the tree will memorize the information from the data, which will not try to learn from the data and will give 100% accuracy in the training set and poor performance in the testing set, finally resulting in [overfitting](#) of the model.

## Q9. Why are Decision Trees Unstable Algorithms?

Decision trees are known as unstable algorithms, as when we add a new data point or new data observation to the existing data, due to the unique data point, the model will need to re-train and add the latest data point in the dataset, here the complete tree structure will now of no use and the new tree be required to include that datapoint in the model. Due to this property of the algorithm, they are known as unstable models.

## Q10. Why is Feature Scaling not Required in Decision Trees?

The decision trees work on splitting the data according to the information gain and entropy from the split. Here the scale of the data is different from the other attributes; it will not affect the entropy and information gain of the split. Hence, it is not necessary to scale of data before applying the said algorithm.

## Conclusion

In this article, we discussed the top interview question that can be asked in data science questions with their appropriate answers. This article will help one to answer these questions efficiently in the interviews and learn a better approach to designing and delivering an accurate and precise response.

Some of the **Key Takeaways** from this article are:

1. Decision trees are tree-based models which splits the data and result output.
2. CART algorithms,s are binary node algorithms, whereas ID3 are multiple node algorithms that can be used for nodes having more than two splits.

3. Very shallow trees result in underfitting, and intense trees result in overfitting.
4. Decision trees are unstable algorithm that trains when a new observation is added.
5. Features scaling is not required for decision trees as it does not affect the information gain and entropy of the split.

**The media shown in this article is not owned by Analytics Vidhya and is used at the Author's discretion.**

---

Article Url - <https://www.analyticsvidhya.com/blog/2023/02/top-10-must-read-interview-questions-on-decision-trees/>



### **Parth Shukla**

UG (PE) @PDEU | 25+ Published Articles on Data Science | Data Science Intern & Freelancer | Amazon ML Summer School '22 | AI/ML/DL Enthusiast | Reach Out @portfolio.parthshukla.live