**Decision Trees (ID3 algorithm)**

**ID3** stands for **Iterative Dichotomiser 3** and is named such because the algorithm iteratively (repeatedly) dichotomizes (divides) features into two groups.

The ID3 algorithm builds decision trees using a top-down greedy search approach through the space of possible branches with no backtracking.

**The steps to construct a ID3 decision tree are as follows:**

1. Calculate entropy for the given dataset.

2. For each attribute/feature:

2.1. Calculate entropy for all its categorical values.

2.2. Take the average information entropy of the current attribute.

2.3. Calculate gain for the current attribute.

3. Pick the highest gain attribute.

4. Repeat until the desired tree is complete.

Is ID3 only binary classification?

**ID3, as an "Iterative Dichotomiser," is for binary classification only.**

**When To Stop Building the Tree?**

When all the leaf nodes are pure (leaf nodes have data that belong to one class)

**Advantages of decision trees**

1. Simple to understand as it follows the same process which a human thinking does in taking decisions in real-life.
2. Useful for solving decision-related problems.
3. Helps to think about all possible outcomes for a problem.
4. There is less requirement of data cleaning compared to other algorithms.

**Disadvantages of decision trees**

1. Contains a lot of layers, which make it complex.
2. For more class labels, computational complexity of decision tree may increase.
3. May have an overfitting issue.

**Overfitting and Decision Trees**

Overfitting can be a big challenge with Decision Trees. We see the algorithm continues to split till it reaches a leaf node. Often the leaf node may just have one or two instances. This will clearly lead to a complex tree structure which may not generalize well to a test scenario. This is because each leaf will represent a very specific set of attribute combinations that are seen in the training data, and the tree will not be able to classify attribute combinations not seen in the training data. There are several ways we can prevent the decision tree from becoming too large. 3 broad approaches to avoiding overfitting are distinguished:

1. Pre pruning or Early stopping: Preventing the tree from growing too big or deep
2. Post Pruning: Allowing a tree to grow to its full depth and then getting rid of various branches based on various criteria

3. Ensembling or using averages of multiple models such as Random Forest.

**Questions on Decision Trees (5 marks each)**

1. What is a decision tree? Explain with the help of an example the working of a decision tree.
2. What is the significance of entropy and information gain in ID3 algorithm?
3. Explain the terms Information Gain and Gini Index.
4. What do you mean by pruning in decision trees.
5. Discuss the advantages and disadvantages of decision tree algorithm.