Decision Tree Classification

Definition - A Decision Tree is a supervised learning algorithm used for classification and regression tasks. In classification, it models data by splitting it into branches to make decisions that lead to a prediction of a target class. It works like a flowchart where each node represents a decision based on one feature, each branch represents the outcome of that decision, and each leaf node represents a class label.

How It Works

1 - **Tree Structure**: Nodes represent features or attributes used for making decisions. Edges represent the outcome of a decision or test, each edge connects a parent node to its child nodes. Branches are the paths from the root to the leaf nodes, showing the decision rules that lead to a particular class label

2 - **Tree Construction**: Tree construction in a Decision Tree involves recursively splitting the dataset based on feature values to create a tree structure. The process starts at the root node, where the dataset is evaluated to determine which feature provides the best split according to a chosen criterion like Gini Impurity or Entropy. This feature is then used to partition the dataset into subsets, each corresponding to a branch of the tree. For each subset, the process is repeated: the best feature is chosen for splitting, and the data is further divided. This recursive splitting continues until the data in each subset is homogeneous (i.e., all instances belong to the same class) or other stopping criteria are met, such as reaching a maximum tree depth or having a minimum number of instances in a node. The result is a tree with nodes representing features, branches representing feature values, and leaf nodes representing the final class labels.

3 - **Making Predictions**: To make predictions, follow the path from the root to a leaf node based on the feature values of the new instance. At each internal node, choose the branch that corresponds to the feature value of the instance. The class label at the leaf node where you end up is the predicted class for the instance.