Decision tree classifiers

Decision tree classifiers work by creating a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. Here's an overview of their operating principle:

1. Tree Structure:
    - A decision tree consists of nodes (decision points) and branches (outcomes).
    - The topmost node is called the root node.
    - Internal nodes represent feature tests.
    - Leaf nodes represent class labels or decisions.
2. Training Process: a. Feature Selection:
    - At each node, the algorithm selects the best feature to split the data.
    - The "best" feature is typically chosen based on metrics like Gini impurity, information gain, or variance reduction.

    b. Splitting:

    - The data is split based on the selected feature's value.
    - For categorical features, splits are typically binary
    - For numerical features, a threshold is chosen

    c. Recursive Partitioning:

    - The process repeats for each resulting subset, creating child nodes.
    - This continues until a stopping criterion is met (e.g., maximum depth, minimum samples per leaf).
3. Prediction:
    - To classify a new instance, it's passed down the tree from the root.
    - At each internal node, a decision is made based on the instance's feature value.
    - The process continues until a leaf node is reached, where the final classification is made.