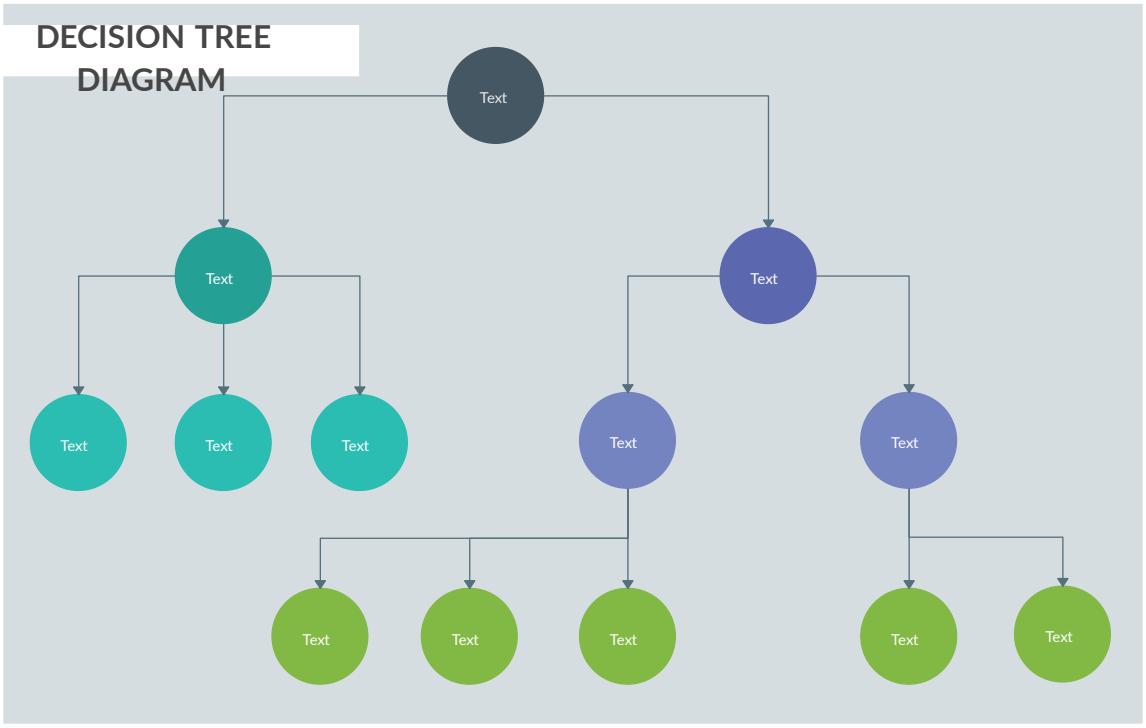
Decision Tree:-

A **Decision Tree** is a supervised machine learning algorithm used for classification and regression tasks .It splits data into branches based on feature values, forming a tree-like structure where each node represents a decision, and leaf nodes represent the final outcome.**Key Components of a Decision Tree**

1. **Root Node**: The topmost node representing the entire dataset.
2. **Decision Nodes**: Nodes where the dataset is split based on a feature.
3. **Leaf Nodes**: Nodes representing the final output (class label or value).
4. **Branches**: The connections between nodes showing decision paths.



How It Works

1. **Select a feature** (using criteria like Gini Impurity, Entropy, or Variance).

2. **Split the dataset** based on that feature.
3. **Repeat recursively** until a stopping condition is met (e.g., max depth, minimum samples per leaf).
4. **Assign a class label or value** at the leaf nodes.

Advantages of Decision Trees

1. Easy to interpret and visualize.
2. Works with both numerical and categorical data.
3. Requires minimal data preprocessing (no need for feature scaling).

Disadvantages

1. Prone to overfitting (mitigated using pruning techniques).
2. Can be biased if dataset is imbalanced.
3. Splitting can be computationally expensive for large datasets.

Applications of Decision Trees

- **Customer segmentation** (e.g., marketing campaigns).
- **Fraud detection** (e.g., financial transactions).
- **Medical diagnosis** (e.g., classifying diseases).
- **Predictive maintenance** (e.g., failure prediction in machinery).