Decision Trees are a type of supervised machine learning algorithm that is used for both classification and regression tasks.
They are known for their simple yet powerful structure that resembles a flowchart, where each internal node corresponds to a feature
(or attribute), each branch represents a decision rule, and each leaf node represents an outcome.

These trees are built by splitting the dataset into subsets based on the value of input features. This process is repeated recursively,
resulting in a tree with nodes that represent decisions based on feature values. One of the most commonly used algorithms to build
decision trees is the CART (Classification and Regression Trees) algorithm, which splits the data using Gini impurity or information gain.

Decision trees are intuitive, easy to interpret, and require little data preprocessing. They can handle both numerical and categorical data,
missing values, and do not assume linear relationships between features. However, they are prone to overfitting, especially if the tree
is very deep and complex. Techniques like pruning, setting a maximum depth, and using ensemble methods such as Random Forests or
Gradient Boosted Trees can help mitigate overfitting.

In practical applications, decision trees are widely used in areas like customer segmentation, credit scoring, medical diagnosis, and
marketing strategies. For instance, a retail company might use a decision tree to predict customer churn based on features such as age,
purchase history, and frequency of store visits.

To illustrate, imagine a decision tree used to determine loan approval. The root node could start by checking the applicant's credit score.

If the score is high, the tree might check income level next. If the score is low, it may look at employment history. The path taken down

the tree represents a series of decisions, ultimately leading to an approval or denial at the leaf node.

Although decision trees are powerful on their own, their true potential is unlocked when used in ensembles. Random Forests combine the

predictions of multiple decision trees to improve accuracy and robustness. Gradient Boosted Trees build trees sequentially, each one

learning from the mistakes of the previous, resulting in a highly effective predictive model.

Overall, decision trees are a vital tool in the machine learning toolbox due to their versatility, interpretability, and effectiveness

across a wide range of problems.