1. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal, but are also a popular tool in machine learning.
2. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal, but are also a popular tool in machine learning.
3. A decision tree is a flowchart-like structure in which each internal node represents a "test" on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules.
4. A decision tree consists of three types of nodes:[1]
   1. Decision nodes – typically represented by squares
   2. Chance nodes – typically represented by circles
   3. End nodes – typically represented by triangles
5. It is one of the most commonly used classification techniques.
6. **Advantages**: Computationally cheap to use, easy for humans to understand learned results, missing values OK, can deal with irrelevant features
7. **Dis-Advantages**: Prone to overfitting
8. **Works with**: Numeric values, nominal values
9. Regression tree analysis is when the predicted outcome can be considered a real number
10. **Types of Decision Trees**

Types of decision tree are based on the type of target variable we have. It can be of two types:

* 1. Categorical Variable Decision Tree: Decision Tree which has categorical target variable then it called as categorical variable decision tree. Example:- In above scenario of student problem, where the target variable was “Student will play cricket or not” i.e. YES or NO.
  2. Continuous Variable Decision Tree: Decision Tree has continuous target variable then it is called as Continuous Variable Decision Tree.

1. **Pruning** - When we remove sub-nodes of a decision node, this process is called pruning. You can say opposite process of splitting.
2. **Both the trees work almost similar to each other, let’s look at the primary differences & similarity between classification and regression trees:**
   1. Regression trees are used when dependent variable is continuous. Classification trees are used when dependent variable is categorical.
   2. In case of regression tree, the value obtained by terminal nodes in the training data is the mean response of observation falling in that region. Thus, if an unseen data observation falls in that region, we’ll make its prediction with mean value.
   3. In case of classification tree, the value (class) obtained by terminal node in the training data is the mode of observations falling in that region. Thus, if an unseen data observation falls in that region, we’ll make its prediction with mode value.
   4. Both the trees divide the predictor space (independent variables) into distinct and non-overlapping regions. For the sake of simplicity, you can think of these regions as high dimensional boxes or boxes.
   5. Both the trees follow a top-down greedy approach known as recursive binary splitting. We call it as ‘top-down’ because it begins from the top of tree when all the observations are available in a single region and successively splits the predictor space into two new branches down the tree. It is known as ‘greedy’ because, the algorithm cares (looks for best variable available) about only the current split, and not about future splits which will lead to a better tree.
   6. This splitting process is continued until a user defined stopping criteria is reached. For example: we can tell the the algorithm to stop once the number of observations per node becomes less than 50.
   7. In both the cases, the splitting process results in fully grown trees until the stopping criteria is reached. But, the fully grown tree is likely to overfit data, leading to poor accuracy on unseen data. This bring ‘pruning’. Pruning is one of the technique used tackle overfitting. We’ll learn more about it in following section.