Decision Tree

INDEED INSPIRING INFOTECH

WWW.INDEEDINSPIRING.COM

KUSHAL@INDEEDINSPIRING.COM

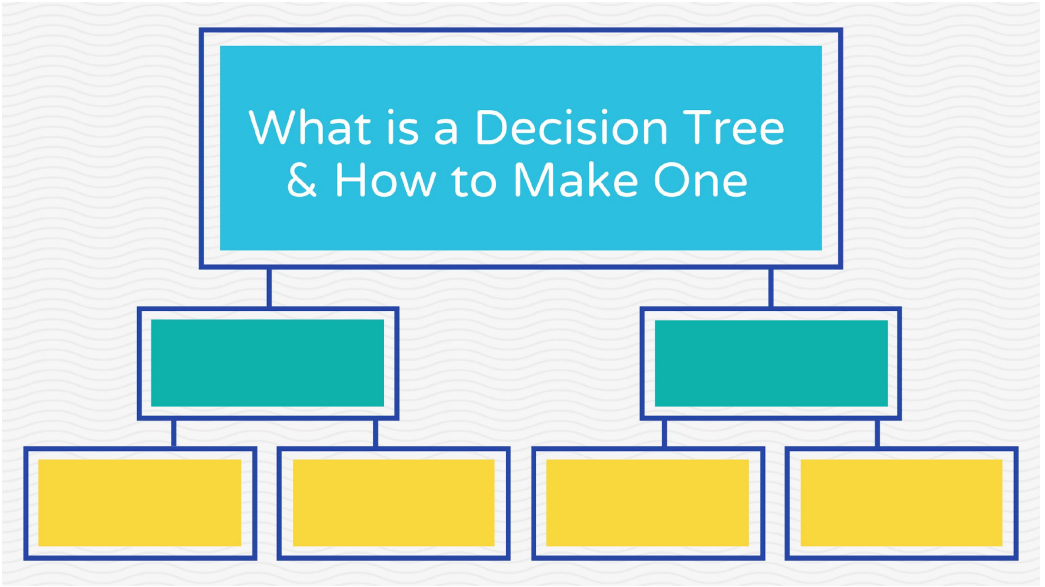
Decision Tree

A decision tree is a graphical representation of specific decision situations that are used when complex branching occurs in a structured decision process.

A decision tree is a predictive model based on a branching series of Boolean tests that use specific facts to make more generalized conclusions.

A Decision Tree is a simple representation for classifying examples.

It is a Supervised Machine Learning where the data is continuously split according to a certain parameter.

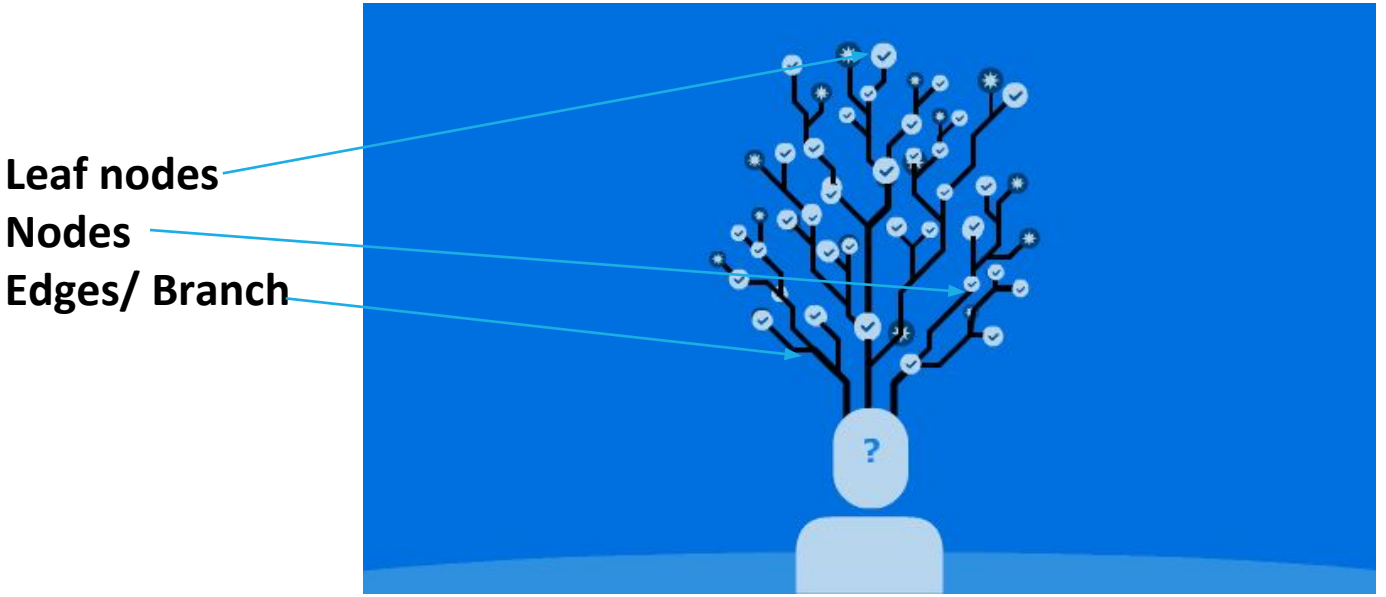


Decision Tree consists of

Nodes : Test for the value of a certain attribute.

Edges/ Branch : Correspond to the outcome of a test and connect to the next node or leaf.

Leaf nodes : Terminal nodes that predict the outcome (represent class labels or class distribution).

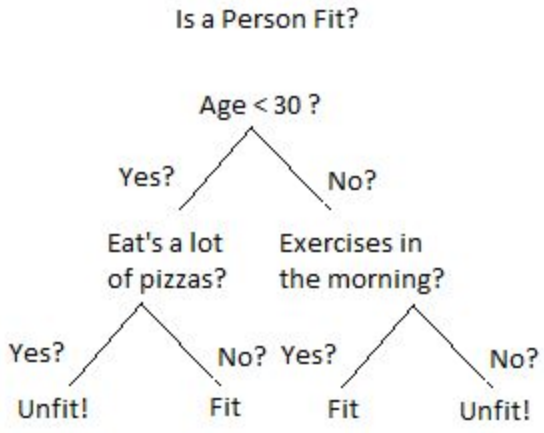


Decision Tree Example

To understand the concept of Decision Tree consider the example.

Let's say you want to predict whether a person is fit or unfit, given their information like age, eating habits, physical activity, etc.

The decision nodes are the questions like 'What's the age?', 'Does he exercise?', 'Does he eat a lot of pizzas'? And the leaves represent outcomes like either 'fit', or 'unfit'.



Types: Classification & Regression

CLASSIFICATION

What we've seen above is an example of classification tree, where the outcome was a variable like 'fit' or 'unfit'. Here the decision variable is Categorical/ discrete.

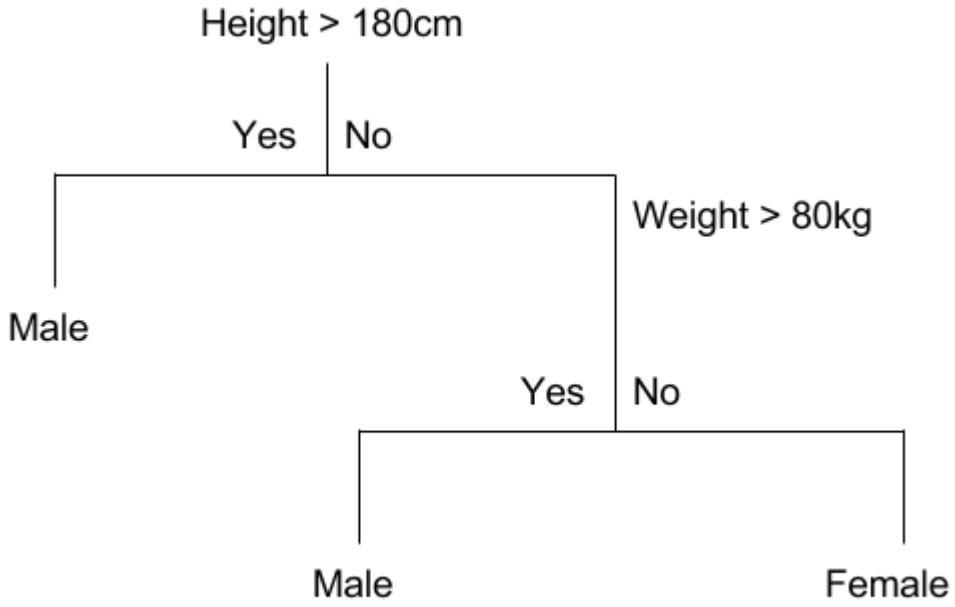
Such a tree is built through a process known as binary recursive partitioning. This is an iterative process of splitting the data into partitions, and then splitting it up further on each of the branches.

REGRESSION

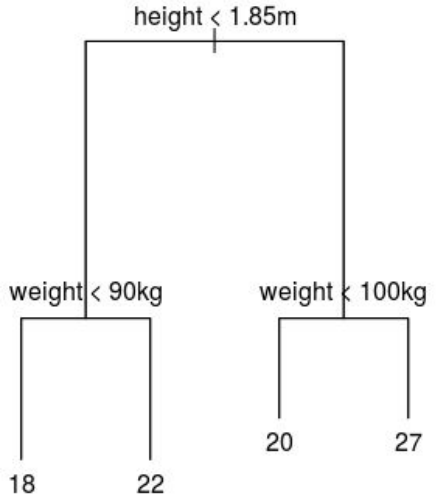
Decision trees where the target variable can take continuous values (typically real numbers) are called regression trees. (e.g. the price of a house, or a patient's length of stay in a hospital)

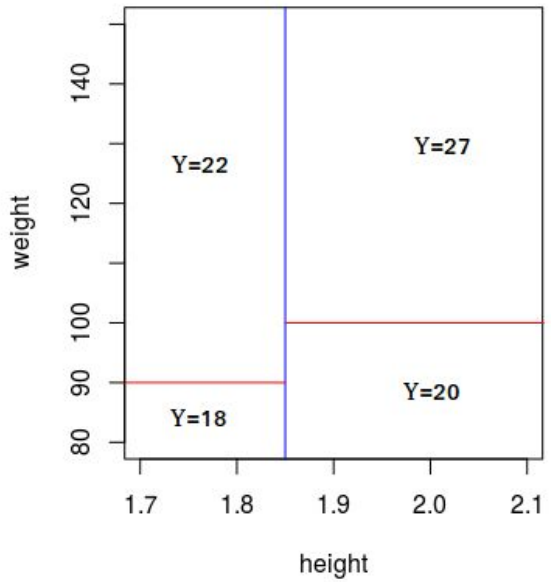
Types: Classification & Regression

CLASSIFICATION



REGRESSION





Applications

To achieve an objective such as maximum profit or optimize cost.

When there are several courses of action.

There is a calculable measure of benefit of various alternatives.

When there are events beyond the control of decision makers: environmental factors.

Uncertainty concerning which outcome will actually happen.

Summary

A decision tree is a graphical representation of specific decision situations that are used when complex branching occurs in a structured decision process.

Decision Tree consists of nodes, edges and leaf nodes.

There are two types of decision trees: classification and regression.