Classification and regression trees is a term used to describe decision tree algorithms that are used for classification and regression learning tasks.

**The Classification and Regression Tree methodology, also known as the CART were introduced in 1984 by Leo Breiman, Jerome Friedman, Richard Olshen, and Charles Stone.**

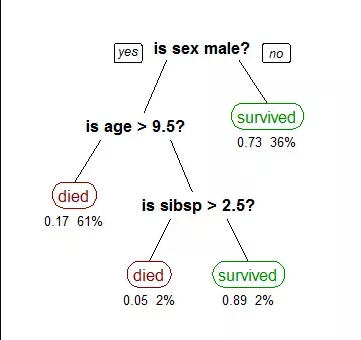
In order to understand classification and regression trees better, we need to first understand decision trees and how they are used.

While there are many classification and regression tree ppts and tutorials around, we need to start with the basics.

**What are Decision Trees?**

**Table of Contents**[[show](https://www.digitalvidya.com/blog/classification-and-regression-trees/)]

If you strip it down to the basics, decision tree algorithms are nothing but if-else statements that can be used to predict a result based on data. For instance, this is a simple decision tree that predicts whether a passenger on the Titanic survived.



Decision Trees

Machine learning algorithms can be classified into two types- supervised and unsupervised. A decision tree is a supervised machine learning algorithm. It has a tree-like structure with its root node at the top.

**Classification and Regression Trees Tutorial**

The CART or Classification & Regression Trees methodology refers to these two types of decision trees.

While there are many classification and regression trees tutorials and classification and regression trees ppts out there, here is a simple definition of the two kinds of decision trees. It also includes classification and regression tree examples.

**(i) Classification Trees**

A classification tree is an algorithm where the target variable is fixed or categorical. The algorithm is then used to identify the “class” within which a target variable would most likely fall.

An example of a classification-type problem would be determining who will or will not subscribe to a digital platform; or who will or will not graduate from high school.

Want to Know the Path to Become a  
**Data Science Expert**?

**Download Detailed Brochure** and Get Complimentary access to Live Online Demo Class with Industry Expert

Date: 18th Dec, 2021 (Saturday) Time: 11:00 AM to 12:00 PM (IST/GMT +5:30)

Top of Form

* 

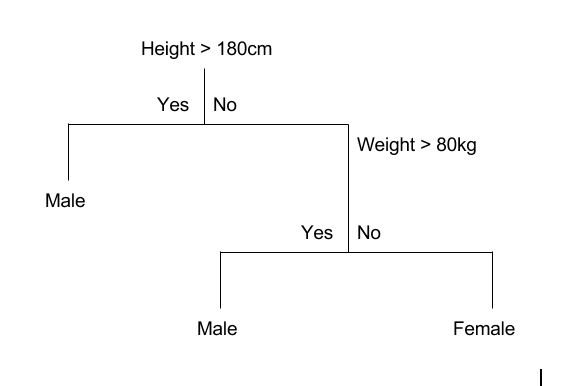


Bottom of Form

By clicking the above button, you agree to our [terms and conditions](https://www.digitalvidya.com/terms-of-use/) and our [privacy policy](https://www.digitalvidya.com/privacy-policy/).

These are examples of simple binary classifications where the categorical dependent variable can assume only one of two, mutually exclusive values. In other cases, you might have to predict among a number of different variables. For instance, you may have to predict which type of smartphone a consumer may decide to purchase.

In such cases, there are multiple values for the categorical dependent variable. Here’s what a classic classification tree looks like.

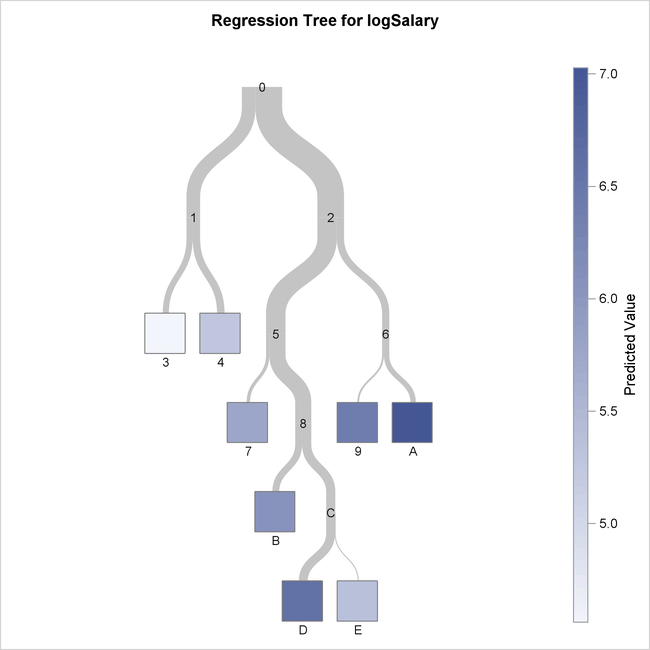


Classification Trees

**(ii) Regression Trees**

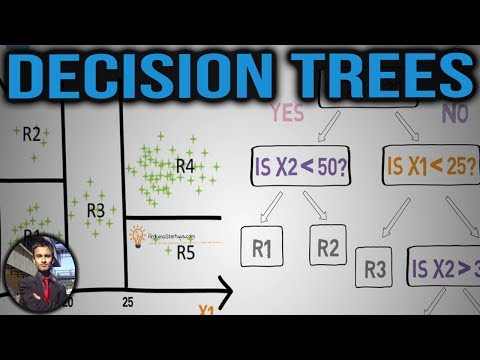
A regression tree refers to an algorithm where the target variable is and the algorithm is used to predict its value. As an example of a regression type problem, you may want to predict the selling prices of a residential house, which is a continuous dependent variable.

This will depend on both continuous factors like square footage as well as categorical factors like the style of home, area in which the property is located, and so on.



Regression Trees

Watch this video for a basic classification and regression trees tutorial as well as some classification and regression trees examples.



**Difference Between Classification and Regression Trees**

[Decision trees](https://www.digitalvidya.com/blog/a-in-depth-decision-tree-learning-tutorial-to-get-you-started/) are easily understood and there are several classification and regression trees ppts to make things even simpler. However, it’s important to understand that there are some fundamental differences between classification and regression trees.

**When to use Classification and Regression Trees**

Classification trees are used when the dataset needs to be split into classes that belong to the response variable. In many cases, the classes Yes or No.

In other words, they are just two and mutually exclusive. In some cases, there may be more than two classes in which case a variant of the classification tree algorithm is used.

Regression trees, on the other hand, are used when the response variable is continuous. For instance, if the response variable is something like the price of a property or the temperature of the day, a regression tree is used.

In other words, regression trees are used for prediction-type problems while classification trees are used for classification-type problems.

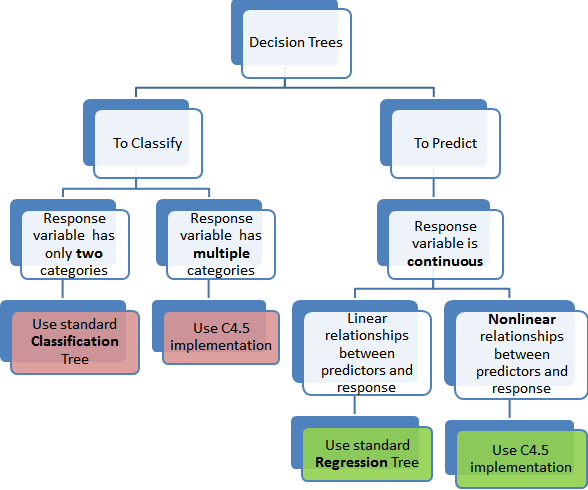
**How Classification and Regression Trees Work**

A classification tree splits the dataset based on the homogeneity of data. Say, for instance, there are two variables; income and age; which determine whether or not a consumer will buy a particular kind of phone.

If the training data shows that 95% of people who are older than 30 bought the phone, the data gets split there and age becomes a top node in the tree. This split makes the data “95% pure”. Measures of impurity like entropy or Gini index are used to quantify the homogeneity of the data when it comes to classification trees.

In a regression tree, a regression model is fit to the target variable using each of the independent variables. After this, the data is split at several points for each independent variable.

At each such point, the error between the predicted values and actual values is squared to get “A Sum of Squared Errors”(SSE). The SSE is compared across the variables and the variable or point which has the lowest SSE is chosen as the split point. This process is continued recursively.



CART Working

**Advantages of Classification and Regression Trees**

The purpose of the analysis conducted by any classification or regression tree is to create a set of if-else conditions that allow for the accurate prediction or classification of a case.

Classification and regression trees work to produce accurate predictions or predicted classifications, based on the set of if-else conditions. They usually have several advantages over regular decision trees.

**(i) The Results are Simplistic**

The interpretation of results summarized in classification or regression trees is usually fairly simple. The simplicity of results helps in the following ways.

* It allows for the rapid classification of new observations. That’s because it is much simpler to evaluate just one or two logical conditions than to compute scores using complex nonlinear equations for each group.
* It can often result in a simpler model which explains why the observations are either classified or predicted in a certain way. For instance, business problems are much easier to explain with if-then statements than with complex nonlinear equations.

**(ii) Classification and Regression Trees are Nonparametric & Nonlinear**

The results from classification and regression trees can be summarized in simplistic if-then conditions. This negates the need for the following implicit assumptions.

* The predictor variables and the dependent variable are linear.
* The predictor variables and the dependent variable follow some specific nonlinear link functions.
* The predictor variables and the dependent variable are monotonic.

Since there is no need for such implicit assumptions, classification and regression tree methods are well suited to data mining. This is because there is very little knowledge or assumptions that can be made beforehand about how the different variables are related.

As a result, classification and regression trees can actually reveal relationships between these variables that would not have been possible using other techniques.

**(iii) Classification and Regression Trees Implicitly Perform Feature Selection**

Feature selection or variable screening is an important part of analytics. When we use decision trees, the top few nodes on which the tree is split are the most important variables within the set. As a result, feature selection gets performed automatically and we don’t need to do it again.