

## Decision Tree Modeling Using R Certification Training

Decision Tree Modelling is a type of supervised learning algorithm and a popular Analytic technique. It can be implemented in various business fields like automobile, telecom or money lending business.

### Objective

### Details

**Duration:** 2 Days

### Who is this course for

- Analytics professionals
- Data Mining professionals
- Graduates who wish to make their career in the analytics industry

Apart from these, the course will also be useful for people working in these areas:

- Data Analysts
- Statisticians
- Data Engineers
- Analytics Consultants
- Information architects
- Developers who want to become a Data Scientist
- Professionals working with SPSS, R, SAS etc

## Course Content

### Introduction to Decision Tree

The module explains Decision Tree, its benefits and core objectives of Decision Tree modelling. The delegates will learn about the gains from the Decision Tree and how it can be applied in business scenarios.

- Decision Tree modeling Objective
- Anatomy of a Decision Tree
- Gains from a decision tree (KS calculations)
- Definitions related to objective segmentations

### Data design for Modelling

The module describes the data design for modelling.

- Historical window
- Performance window
- Decide performance window horizon using Vintage analysis
- General precautions related to data design

### Data treatment before Modelling

In this module, the delegates will learn about data sanity check and how to perform the necessary checks before modelling.

- Data sanity check-Contents
- View
- Frequency Distribution
- Means / Uni-variate
- Categorical variable treatment
- Missing value treatment guideline
- Capping guideline

### Classification of Tree development and Algorithm details

The module explains how to develop the decision tree using R and the Algorithm.

- Preamble to data
- Installing R package and R studio
- Developing first Decision Tree in R studio
- Find strength of the model
- Algorithm behind Decision Tree
- How is a Decision Tree developed?
- First on Categorical dependent variable
- GINI Method
- Steps taken by software programs to learn the classification (develop the tree)

### Industry practice of Classification tree - Development, Validation and Usage

The module describes the development and validation of Classification trees.

- Discussion on project
- Find Strength of the model
- Steps taken by the software program to implement the learning on unseen data
- Learning more from a practical point of view
- Model Validation and Deployment

## Regression Tree and Auto Pruning

In this module, the delegates will learn about the Advance stopping criteria of a decision tree.

- Introduction to Pruning
- Steps of Pruning
- Logic of pruning
- Understand K fold validation for model
- Implement Auto Pruning using R
- Develop Regression Tree
- Interpret the output
- How it is different from Linear Regression
- Advantages and Disadvantages over Linear Regression
- Another Regression Tree using R

## CHAID Algorithm

The module explains Chi square and CHAID and difference between CHAID and CART etc.

- Key features of CART
- Chi square statistics
- Implement Chi square for decision tree development
- Syntax for CHAID using R, and CHAID vs CART

## Other Algorithms

In this module, the delegates will learn about Random Forest, ID3 and Entropy.

- Entropy in the context of decision tree
- ID3
- Random Forest Method
- Using R for Random forest method