Tel +44 20 4571 2395
info@siliconbeachtraining.co.uk

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. The Decision Tree Modeling Using R Certification training course provides an overview of the decision tree and algorithm behind decision tree, i.e. how does decision tree software work. The course equips the delegates with the necessary skills required to become a Predictive analytics expert. The delegates will gain a crystal clear understanding of Decision Tree model, benefits of learning the technique, how to apply the decision tree techniques and perform the validation.

The delegates will also be introduced to the advanced concepts like CHAID, CART, Regression Tree, Pruning and data designing through the training program. The two day course covers the benefits of modeling, predictive analytics, objective segmentation and how to interpret and implement the Decision tree model. During the Decision Tree course, the delegates will learn about decision tree, the various algorithm behind decision tree, how to interpret the decision tree output of R, what benefit it brings and steps involved in developing decision tree in R.

Prerequisites

The delegates must have basic knowledge of R programming language before attending the course. The course throws light on R programming syntax required for the development of Decision Tree model.

Course Objectives

- Gain a thorough understanding of a Decision Tree
- Learn how to implement various Decision Tree techniques (CHAID / CART etc.)
- Learn where to use CHAID / CART / ID3 etc.
- Learn to use the R platform to develop Decision Trees
- Learn to design data for Decision Tree modeling
- Implement Decision Trees to derive business insights
- Perform Decision Tree Model Validation
- Interpret and Implement Decision Tree model

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

Tel +44 20 4571 2395
info@siliconbeachtraining.co.uk

- 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.

Tel +44 20 4571 2395
info@siliconbeachtraining.co.uk

- 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

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.