ThinkPod Academy

A unique place to learn, network and find your dream job

Course Demo

Week 1: *Basics of R Programming and Statistics*

1. **Introduction to R:**
   a. *Installation of R Studio, implementing simple mathematical operations and logic using R operators, loops, if statements*
   b. *Data Exploration: Importing and exporting data from/to external sources, working with data frames, in-built functions, matrix, list and array*
   c. *Data Manipulation : Subset Datasets, data transformation, Sampling , Summarizing and SQL in R*
   d. *Data Visualization: Introduction to ggplot2 package.*

2. **Introduction to Statistics:**
   a. *Measures of central Tendency*
   b. *Measures of dispersion*
   c. *Types of Variables, Introduction to Probability and density functions*
   d. *Central Limit Theorem*
   e. *Hypothesis Testing (Null Hypothesis, Z test, T test, P value, ANOVA, Type1 & Type 2 errors, chi-square test)*
   f. *Common Probability Distributions (Normal, Standard Normal, T distribution, Poisson, Chi Square)*

Week 2: *Traditional Modelling Techniques using R*

1. **Supervised vs Unsupervised Algorithms**

2. **Linear Regression:**
   a. *Use cases*
   b. *Predictive modelling vs deterministic modelling*
   c. *Population vs Sample based modeling*
   d. *OLS technique and Cost Function in Linear Regression*
   e. *Assumptions of Linear Regression and violations of assumptions*
   f. *Variable Reduction Techniques (Theory and in R)*
   g. *Performance metrics used in Linear Regression*
   h. *Linear regression in R and interpreting the results*

3. **Logistic Regression:**
   a. *Use cases*
   b. *Introduction to binary/ categorical outcomes and Why not OLS when outcome is binary*
   c. *Introduction to MLE and Cost Function in Logistic Regression*
   d. *Variable Reduction Techniques in Theory and in R (Information Value, Multicollinearity, Stepwise Selection etc.)*
   e. *Performance metrics used in Logistic Regression*
   f. *Logistic Regression in R and interpreting the results*

Week 3: *Machine Learning Techniques using R*

1. **Segmentation:**
   a. *Use cases*
   b. *Types of segmentation*
   c. *Introduction to Decision Trees*
      i. *CHAID*
      ii. *CART*
      iii. *Information Gain, Entropy, Gini Index*
      iv. *Decision Trees in R*

2. **Random Forest and GBM:**
   a. *Bias Variance Tradeoff*
   b. *Weak Learners Concept*
   c. *Types of Validation*
   d. *Concept of Bootstrap Aggregation*
   e. *Random Forest Algorithm and Prediction methods*
   f. *Concept of Boosting*
   g. *GBM Algorithm and Prediction methods*
   h. *RF and GBM in R*

Week 4

1. **Case Study in R**
   a. *Understand and use the techniques from first 3 weeks to solve a real-time problem*
2. **Resume Preparation**
3. **Mock Interview**
4. **Q&A**