



# BUSINESS ANALYTICS AND DATA MINING MODELING USING R

## PROF. GAURAV DIXIT

Department of Management Studies  
IIT Roorkee

**PRE-REQUISITES** : Basic Statistics Knowledge

**INDUSTRIES APPLICABLE TO** : Big Data companies, Analytics & Consultancy companies, Companies with Analytics Division

## COURSE OUTLINE :

Objective of this course is to impart knowledge on use of data mining techniques for deriving business intelligence to achieve organizational goals. Use of R (statistical computing software) to build, assess, and compare models based on real datasets and cases with an easy-to-follow learning curve.

## ABOUT INSTRUCTOR :

Prof. Gaurav Dixit is an Assistant Professor in the Department of Management Studies at the Indian Institute of Technology Roorkee. He earned his doctoral degree from the Indian Institute of Management Indore and an engineering degree from Indian Institute of Technology (BHU) Varanasi. Previously, he worked in Hewlett-Packard (HP) as software engineer, and Sharda Group of Institutions as project manager on deputation. Gaurav's research focuses on information technology (IT) strategy, electronic commerce, electronic waste, data mining, and big data analytics and provides insights on business and social value of IT. His research has appeared in quality journals & conferences, including Resources, Conservation and Recycling, Journal of Global Information Technology Management, Sustainable Production and Consumption, Journal of Information Technology Management, DIGITS conference, India Finance Conference, Indian Academy of Management conference, and Academy of Management conference.

## COURSE PLAN :

**Week 1:** General Overview of Data Mining and its Components Introduction and Data Mining Process Introduction to R Basic Statistical Techniques

**Week 2:** Data Preparation and Exploration Visualization Techniques

**Week 3:** Data Preparation and Exploration Visualization Techniques Dimension Reduction Techniques Principal Component Analysis

**Week 4:** Performance Metrics and Assessment Performance Metrics for Prediction and Classification

**Week 5:** Supervised Learning Methods Multiple Linear Regression

**Week 6:** Supervised Learning Methods Multiple Linear Regression

**Week 7:** Supervised Learning Methods Naïve Bayes

**Week 8:** Supervised Learning Methods Classification & Regression Trees

**Week 9:** Supervised Learning Methods Classification & Regression Trees

**Week 10:** Supervised Learning Methods Logistic Regression

**Week 11:** Supervised Learning Methods Logistic Regression Artificial Neural Networks

**Week 12:** Supervised Learning Methods and Wrap Up Artificial Neural Networks Discriminant Analysis Conclusion.