

Big Data Analytics-1



Open Programme:

May 8-10

Prof. Pulak Ghosh (Presidency Univ)

Data analytics is the process of examining large amounts data of a variety types to uncover hidden patterns, unknown correlations and other useful information to facilitate applications in the areas of customer analytics, banking, retail etc. Successful organizations are focused on using data analytics techniques that drive the greatest business value. Data analytics refers to the ever increasing volume, velocity, variety, variability and complexity in the collected data.

Instructor

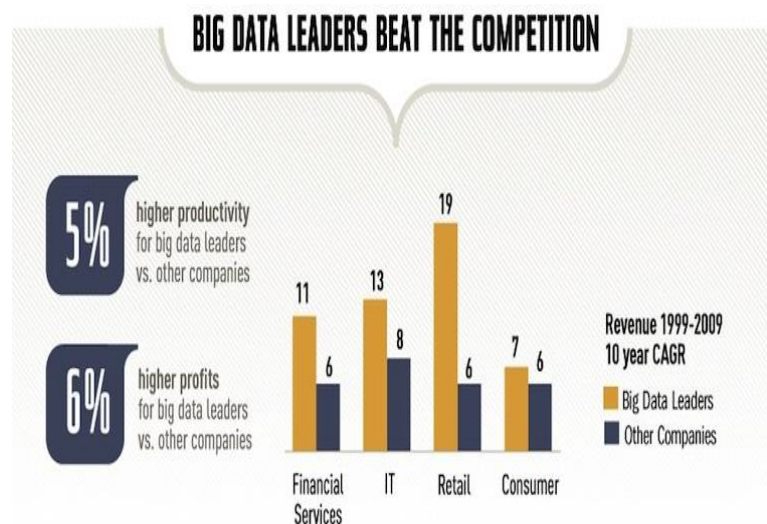
: Prof. Pulak Ghosh, Presidency
University

Prof. Bhuvanesh Pareek, IIM Indore

COURSE OVERVIEW

Big data refers to the ever increasing volume, velocity, variety, variability and complexity in the collected data. Most MBA students who go into management consulting or marketing positions will find themselves having to advise their firm on critical aspects related to product management, understanding customer behavior, and improving the performance of a product in the market. In today's world all these need looking and analyzing large data, called Big data. Big data refers to the ever increasing volume, velocity, variety, variability and complexity.

For organizations, big data is the fundamental consequence of the new marketing landscape, born from digital world we now live in. Think about the customer data collected today—online purchase data, click through rates, browsing behaviour, social media interactions, mobile device usage, geo-location etc. By combining expertise of big data with an integrated marketing analytics can make a substantial impact and reveal richer insights.



Source: McKinsey & Company

Good analytics can make all the difference for effective campaigns and company growth, according to data from McKinsey on Marketing and Sales. According to McKinsey's statistics, big data leaders have, on average, 5% higher productivity and 6% higher profits than other companies --*McKinsey & Company*

Syllabus Outline and Topics covered

This three day workshop systematically covers the various stages of big data and analytics.

Topic-1: Introduction and Overview on analytics: Supervised and Unsupervised Learning

Will learn how machine learning transform the modeling and data understanding in analytics that transform organizational decisions, Making, pitfalls, and payoffs of the business and the three pillars of Analytics: descriptive, predictive, and prescriptive. We will concentrate on plenty of examples in this topic to get the audience conversant on the utility of data analytics.

Readings:

1. HBR Case Study: The Big Idea: The Next Scientific Revolution
2. Competing on Analytics – HBR article (2006)
3. HBR article: 10 Insights: A First Look at The New Intelligent Enterprise Survey on Winning With Data & 10 Data Points: Information and Analytics at Work

Topic-2: Data Science Algorithm: Introduction to R, Python and Knime

Basic introduction to R, Python, Knime, manipulation of data, working with data from various sources, data visualization. Doing analytics in R, Python and Knime.

Descriptive statistics, Bayes theorem and applications to decision making

We will just get introduced to R so that when instructor does some example in r, the audience can relate to it.

Topic 3: King of Prediction- Regression

Why regression is so important?

Regression model building framework: problem definition, Data pre-processing; model building; diagnostics and validation

Simple linear Regression: coefficient of determination, significance tests, Residual analysis, confidence interval

Multiple linear regression: Interpretation of regression coefficients, categorical variables, heteroscedasticity, Multi-collinearity

High Dimensional Regression

Advanced Regression for Prediction

Case studies:

Pedigree vs. Grit: Predicting Mutual fund Manager Performance (Kellogg case)

Topic 4: Classification Methods

Decision trees

KNN clustering

Logistic Regression

Multinomial Logistic Regression

Case Studies:

Several Cases will be discussed

Profile of Program Director

Prof Pulak Ghosh: Prof. Ghosh is currently Infosys chair professor of Economics at Presidency University, Kolkata. He works in the domain of Big data, Machine learning, Marketing analytics, Business analytics, Banking analytics, Econometrics and Bayesian Statistics.

Prof. Ghosh is a member of the Data Privacy Advisory Group of Global Pulse- the UN Secretary General's Big Data Initiative. Ghosh, the only expert from India, along with international experts, will advocate responsible use of Big Data for sustainable development and humanitarian value creation. Prof. Ghosh is also an academic fellow at the Center for Advanced Financial Research and Learning of Reserve Bank of India, He is recently named as the top 10 most influential analytics leader.

Benefits of Taking this Course

- Finding patterns in your data.
 - Provide insights as to what will happen next.
 - Better alignment with strategy.
 - Gain new knowledge and insights.
 - Learn from past and current data to better.
 - Understand future trends.
 - Faster Decision-making.
-