

PRE-CONVENTION TUTORIAL ON DATA SCIENCE



Venue: PSG College of Technology
Date: 01.22.2017

PRE-CONVENTION TUTORIAL ON DATA SCIENCE

SESSION 1

INTRODUCTION TO DATA SCIENCE (Speaker: Karthik Ramasubramanian):

Data Science – Three Pillars of Data Science (Hacking Skills, Math & Statistical Knowledge and Substantive experience) – Data Structure & Machine Learning – Big Data – Visualization – Cloud – Internet of Things – Statistics – Enlighten the facts and various fields where the Data Science is applied. Educated the problem solving mechanism using data science and predefined algorithms.

CASE STUDY ON BMI CALCULATION (Speaker: Abishek Singh)

Challenges in insurance – Process of getting life insurance – Prediction of BMI.

STATISTICS

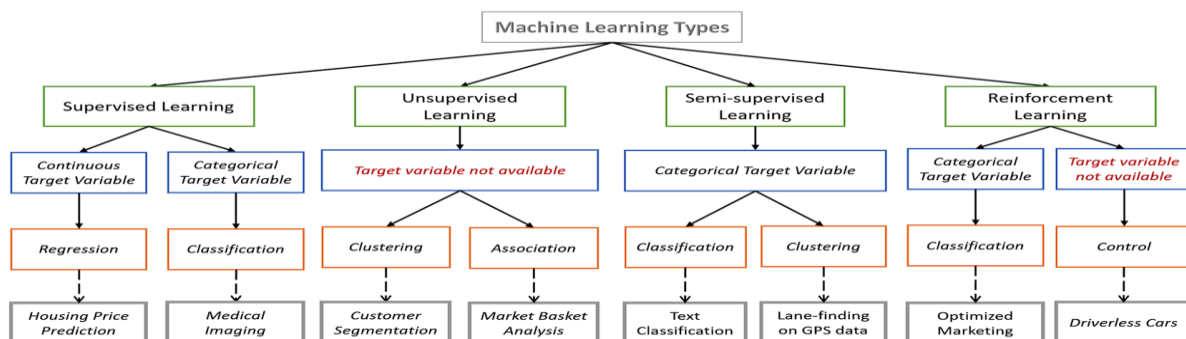
Data descriptive: Structured, Semi Structured, Quasi structured, Unstructured data.

Machine Learning (ML): ML is an algorithm that can learn this relationship without relying on any rule-based programming. ML will emphasize on how the final predictions will look like if similar data is supplied in future.

Statistical Learning: Statistical modeling will estimate the relationship based on formal quantification from statistical inferences. The process of statistical inference quantifies the process by which data is generated.

SESSION 2- (Speaker: Abishek Singh)

Machine Learning Algorithm – K-means, Decision Tree & Neural Network.



Visualization – GGPlots in R – Boxplot, Histogram, Scatterplot, Sankey Plot, Cohort Charts, Bubble Chart. GGviz – Motion Charts.

Ensembles: Bias vs Variance trade off - Bagging & Booting.

SESSION 3- (Speaker: Karthik Ramasubramanian):

Big Data – Ecosystem –



Graph Databases - Graph databases support a very flexible and fine-grained data Model. RDBMS provides results, Graph Databases provides answers.

SESSION 4 - (Speaker: Abishek Singh)

$$\text{BMI} = \begin{cases} \frac{\text{weight}(\text{kg})}{\text{height}(\text{m})^2} \\ \text{or} \\ \frac{\text{weight}(\text{lb}) \times 703}{\text{height}(\text{in})^2} \end{cases}$$

BMI Calculation -

Image Processing – Face Detection – Facial key Points detection.

Use case: Predict Emotion, Recognition – verification.

Industry – Academia Collaboration: Fake news detection, Persona identification, Computer vision & Emotion Detection.

Future of Data Science: Deep Learning – Internet of Things – Online Machine Learning (Mobile Devices & Real –Time Performance).

Talent Graph: Technology Stack - Neo4j, Titan, Python, Cloud APIs & MEAN technology stack

Machine Learning Probabilistic Graph Models, Regression Analysis, Natural Language Processing(NLP).

