NOC: Introduction to Data analytics - Video course

COURSE OUTLINE

Data Analytics is the science of analyzing data to convert information to useful knowledge. This knowledge could help us understand our world better, and in many contexts enable us to make better decisions. While this is broad and grand objective, the last 20 years has seen steeply decreasing costs to gather, store, and process data, creating an even stronger motivation for the use of empirical approaches to problem solving. This course seeks to present you with a wide range of data analytic techniques and is structured around the broad contours of the different types of data analytics, namely, descriptive, inferential, predictive, and prescriptive analytics.

COURSE DETAIL

<table>
<thead>
<tr>
<th>Week No</th>
<th>Topics</th>
</tr>
</thead>
</table>
| 1.      | Descriptive Statistics  
Introduction to the course  
Descriptive Statistics  
Probability Distributions |
| 2.      | Inferential Statistics  
Inferential Statistics through hypothesis tests  
Permutation & Randomization Test |
| 3.      | Regression & ANOVA  
Regression  
ANOVA(Analysis of Variance) |
Differentiating algorithmic and model based frameworks  
Regression: Ordinary Least Squares, Ridge Regression, Lasso Regression, K Nearest Neighbours Regression & Classification |
| 5.      | Supervised Learning with Regression and Classification techniques -1  
Bias-Variance Dichotomy |

Pre-requisites:

This course requires that you are familiar with high-school level linear algebra, and calculus. Knowledge of probability theory, statistics, and programming is desirable.

Coordinators:

- Dr. Balaraman Ravindran  
Department of Computer Science and Engineering\ IT Madras
- Dr. Nandan Sudarsanam  
Department of Management Studies\ IT Madras
<table>
<thead>
<tr>
<th>Model Validation Approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
</tr>
<tr>
<td>Linear Discriminant Analysis</td>
</tr>
<tr>
<td>Quadratic Discriminant Analysis</td>
</tr>
<tr>
<td>Regression and Classification Trees</td>
</tr>
<tr>
<td>Support Vector Machines</td>
</tr>
</tbody>
</table>

6. **Supervised Learning with Regression and Classification techniques -2**
   - Ensemble Methods: Random Forest
   - Neural Networks
   - Deep learning

7. **Unsupervised Learning and Challenges for Big Data Analytics**
   - Clustering
   - Associative Rule Mining
   - Challenges for big data analytics

8. **Prescriptive analytics**
   - Creating data for analytics through designed experiments
   - Creating data for analytics through Active learning
   - Creating data for analytics through Reinforcement learning

**References:**
