

Program Name: Data Science and Machine Learning Overview			
	Session	Objectives	Program Structure
Day 1	Introduction To Data Science <ol style="list-style-type: none"> 1. What is Data Science ? 2. Data Science team structure 3. Data Science stages 4. Machine Learning and data science 	Introduction to data science discipline as an approach to extract hidden patterns from data Skills required in Data Science, Structure of data science team	Presentations, discussions
	Introduction to Machine Learning <ol style="list-style-type: none"> 1. What is Machine Learning ? 2. Why Machine Learning 3. Requisites for Machine Learning 	Introduce participants to machine learning concept Applications of machine learning with examples Pre-requisites for machine learning	Presentations, discussions, Case studies on implementation of machine learning
	Preparing for ML projects <ol style="list-style-type: none"> 1. Defining the objectives 2. Identifying the required data items 3. Identifying sources of data 4. Data cleansing 5. Preparing data for ML 	Introduce the approach to machine learning project with focus on clarity of objectives, identifying sources of data, preparing data for analytics Hadoop stack and its applications	Presentations, discussions, hands on?
	Data Patterns, Statistical Modelling <ol style="list-style-type: none"> 1. Patterns in data, what does it mean? 2. Representing reality in models 3. Supervised machine learning 4. Unsupervised machine learning 5. Challenges of machine learning 6. Generalization and model fit 	To help participants understand what patterns in data mean To familiarize participants with the two broad classification of machine learning styles, their applicability, requirements of each type Concepts of over fitting / under fitting and generalization	Presentations and discussion on live case studies
Day 1	Variables, Attributes and Relations <ol style="list-style-type: none"> 1. What are variables ? 2. Dependent, independent variables 3. Scales 4. Importance of relation between variables 5. Concept of significance in analyzing variable relations 6. Concept of normal distribution 7. Statistical reasoning 	Introduction to types of variables and how variables influence each other. The importance of identifying correct variables that link to the objectives	Presentations, discussions
	Descriptive Statistics <ol style="list-style-type: none"> 1. Central values 2. Measure of variance 3. Shape of distribution 4. Correlations 5. Regression 	A thorough introduction to metrics about data such as mean, median, quantiles, mode, relation between variables	Presentations, discussions, Case studies on implementation of machine learning
Day 2	Inferential Statistics <ol style="list-style-type: none"> 1. Concept of samples and population 2. Hypothesis and hypothesis testing 3. Infer about population from sample 4. Confidence levels, intervals and P Value, Z scores 	A conceptual introduction to inferential statistics	Presentations, discussions, hands on? This will not be an in-depth mathematical session

Day 2- 3	Supervised Learning Methods <ol style="list-style-type: none"> 1. Linear regression 2. Decision trees 3. Naive Bayesian classifiers 	Introduce participants to supervised learning approach with focus on model generation through training data, testing the model , interpreting the results	Presentations, discussions and hands-on coding to implement a POC
	Unsupervised Learning Methods <ol style="list-style-type: none"> 1. Clustering 	To explain concepts of unsupervised machine learning, their applications and how systems learn on their own	Presentations, discussions and demos
	Machine Learning Project (for self learning) To be decided	To explore the avenues for machine learning in current context	Student presentations