Data Science

Description:
Data science is a multi-disciplinary field that uses scientific method, process, algorithms and systems to extract knowledge and insights from structured and unstructured data. Data science is the same concept as data mining and big data: “use the most powerful hardware, the most powerful programming systems, and the most efficient algorithms to solve the problems”. Data science is a “concept to unify statistics, data analysis, machine learning and their related methods” in order to “understand and analyze actual phenomena” with data. It employs techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, and information science.

Day 1:
- What is Data and Data Science?
  - Introduction to Data Science
  - Understanding the need
  - Understanding Big data and machine learning
  - Running machine learning under Linux platform
  - Platforms of working with
Data Science  ▶  Programming
Languages used in Data Science  ▶  Role of Python and R programming in this domain
▶ Basic Introduction of Python syntax and programming logics
▶ Applications of Data Science

Day 2:

- Python programming
  ▶ Basic of python and why python for machine learning
  ▶ Installation of software on different OS.
  ▶ Understanding basic syntax with data types
  ▶ Number, string, list, tuple and dictionary
  ▶ Loops, conditions
  ▶ User input and user defined functions
  ▶ How to use libraries
  ▶ Creating and importing own library

- Project:- Designing of ATM Prototype

- Installation of Python Libraries in System Numpy
  ▶ Formation of Arrays and its operation
  ▶ Working with multidimensional arrays using
numpy
- Data formation and matrix manipulations
- Use of Numpy in Data Science Matplotlib
- Data Visualisation in linear graphs
- Bar Graph
- Multiple data visualisation in one scatter plot

• Data Science Applied to Machine Learning
  1) Supervised and Unsupervised Learning
  2) Working with Python for ML
     - Types of learning
     - Supervised Learning lab with Hello World Program
     - Classification and regression
     - Training your machine with real time datasets

• Project:- Creating own ML datasets and it’s implementation

#) Pandas
- Introduction to dataframes
- Reading / writing data files
- Structure of dataframes
- Use of dataframes with ML
- Datasets reading for Scikit-Learn

Project:- Accessing real datasets using Pandas

**Day 3:**
• Image Processing and Its Application using DS OpenCV
  ■ Image Processing with Python
  ■ Image Read and type conversions
  ■ Image Display in various modes
  ■ Live Image Processing
  ■ Camera Detection and Image Capturing
  ■ Keyboard Interruptions to the code
  ■ Live Image Processing Music Player Libraries with Python
  ■ Download and importing library into python
  ■ Different modes of operations and their control

Wikipedia an Open Source Library
  ■ Searching information on wikipedia
  ■ Wikipedia with ML

• Project:- Customised Music Player using Python
• Project:- Realtime Face Detection System based Music Player