Programming for ML and Data Science

Course Outline (6 Weeks)

Week 1: Introduction to Python Programming

- Introduction to Python and the programming environment
- Variables, Types, Operations
- Tools' installation and execution of a few programs that introduce basic Python concepts;

Week 2: Python for ML and Data Science

- Program flow and control
- Important libraries: numpy, scipy, matplotlib
- Introduction to other important libraries – sklearn, statsmodels

Week 3: ML Overview, Regression

- Measurement Levels in Statistics (Nominal, Ordinal, Interval and Ratio)
- What is ML - Overview
- Types of problems solved in ML (link with Measurement Levels)
    - Supervised, Unsupervised
    - Regression, Classification, Clustering
- Regression: Introduction and applications
- Regression Metrics and interpreting results

Week 4: Regression using python

- Regression using Python
    - Introduction to sklearn
    - Introduction to statsmodels and OLS
    - Interpretation of results
- Other Regression methods
- Invoking multiple Regression methods, followed by selection

Week 5: Classification using Python

- Classification – Introduction and applications
- Logistic Regression
- Other Classification methods

Week 6: Clustering using Python

- Clustering: Introduction and applications
- KMeans and Hierarchical Clustering
- Assessing the quality of clusters: Silhouette Score and Davies-Bouldin Index, etc.
- Clustering using sklearn: Examples