Machine Learning using Python

AI and Machine Learning is emerging as a hot new profession and academic discipline. Harvard Business Review says Data Scientist is the Sexiest Job of the 21st Century. But demand for data scientists is racing ahead of supply. People with the necessary skills are scarce, primarily because the discipline is so new. This course is designed to give a start and introduction to this new discipline. This course is spread across 2 days and will have a plenty of hands on exercises on real world data.

Pre-requisites:

- ✓ A basic understanding of data and programming is required.
- ✓ Programming knowledge using Python is essential

Topics:

- Understanding the Data Science landscape and problem domain
- Accessing, preparing and exploring data with Pandas & Scipy
- EDA: Univariate, Bivariate and Multivariate Analysis
- Feature engineering, missing values imputation, scaling and categorical encoding
- Regression, classification algorithms
- Linear Regression, KNN, Decision Tree, Logistic Regression, Random Forest
- Feature Engineering
- Clustering and Dimensionality Reduction
- Building and validating models
- Evaluation metrics
- Model Deployment

Duration: 24 Hours

Machine Learning using Python

Course Outline

Machine Learning Overview	Introduction to Machine Learning Setting up Python Environment for Data Analysis Overview of Python Stack of Machine Learning - Numpy, Pandas
Exploratory Data Analysis	Basic Statistical analysis using scipy.stats Drawing Histograms, Bar charts, Density Plots, Box Plots Drawing Density plots and understating data distributions Univariate Analysis, Bivariate and Multivariate Analysis
Supervised Learning: Regression	Feature Engineering: Scaling, Categorical Encoding, Imputation Supervised Learning: Regression Regression: Linear Regression, KNN Model Evaluation: K-Fold Cross Validation Metrics: RMSE, R2 score Creating Model Pipeline All Hands on using Scikit-learn library
Supervised Learning: Classification	Feature Engineering: Scaling, Categorical Encoding, Imputation Supervised Learning: Classification KNN, Decision Trees, Ensemble Methods: Random Forest, Boosting Model Evaluation: Confusion Matrix, Precision, Recall
Unsupervised Learning: Clustering and Dimensionality Reduction	Clustering: K-Means Finding optimal Clusters: Elbow Method, Cluster Map Dimensionality Reduction: PCA (Principal Component Analysis)
Model Validation and Deployment	Cross validations Hyperparameter Tuning Model Deployment