

Python™ Predictive



Length: 2 Days / 4 Half Days

Overview: This course is designed to teach and to practice the fundamentals of using Python as a Data Analysis tool for making predictions.

Objectives: The objective of this course is to practice Data Science Analytics and Machine Learning within the context of Predictive analytics, making predictions about the future with data.

Topics include:

- Machine Learning (ML)
- Classification
- Regression
- Clustering
- Supervised vs. Unsupervised Learning
- Charting Review
- Pandas
- NumPy Arrays
- Grouping
- Data Cleaning for ML
- Scaling Parameters
- Parameter Tuning
- Aggregations

Data Exploration Review

- Basic Plotting
- Grouping
- Data Diagnosis Review
- Making Predictions without ML

- Changing Threshold
- Score, Accuracy Score, Cross validation, ROC, AUC, Precision, MSE, and Recall
- Classification Report

Machine Learning Models

- Linear Regression
- Logistic Regression (Classification)
- Properly Preparing Data for ML
 - Data Cleaning
 - Nulls
 - Data Relevance
 - Converting Categorical Data to Numerical Data
 - Dummies
 - SKLearn (Label Encoder, SimpleImputer)
 - Pandas Categorical Data Type
 - Brute Force
 - Normalization and Scaling
 - Correlation

Testing Various Models

- Choosing the best estimator
- DecisionTreeClassifier
- RandomForestClassifier

Grid Search and Manual Parameter Tuning

- Setting up Parameter Grid
- Plotting to track accuracy
- Tracking training and testing accuracy

Practice Perfecting Regression Models

Other General Useful Techniques:

- Sentiment Analysis
- XGBoost
- Prescriptive Analysis
 - Clustering
 - “The Elbow Method”
 - Applying Clusters Practically
 - Using PCA(Principle Component Axis)

Perfecting Models

- Checking Accuracy & Scoring
 - Relative Frequency in Unique Values
 - Prediction and Probability Prediction



Prerequisites

Python Data Prep

Materials

- All students will receive slides with lecture material and data and labs.
- Software Needed on Each Student PC
 - Microsoft Excel 2010 or later (2013 or later recommended)
 - Internet access
 - Jupyter Notebook
- Related data and lab files will be provided

Other Python Courses Available: Python Data Prep; Python II Diagnostic