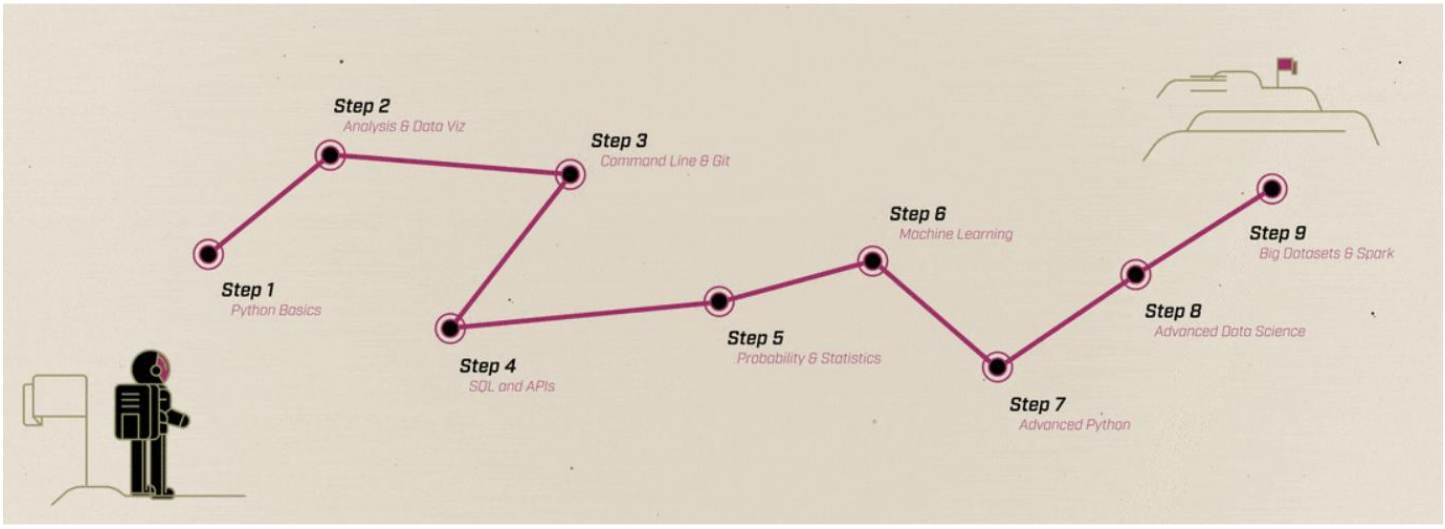
Machine Learning using Python Course outline



Python Basics:

Introduction:

- What is programming?
- What is not programming?
- What kind of things programmer can build?

Setting up our programming environment:

- Programming Environment
    - Configuring Ubuntu for Python 3.3

Python Basics:

- Hello World
- Variables, strings, and Numbers
- Lists and Tuples
- Introducing Functions
- If Statements
- While Loops and Input
- Basic Terminal Apps
- Dictionaries
- More Functions
- Classes

- Exceptions
- Testing

Python For Data Analysis:

Numpy:

- Introduction to numpy
- Creating arrays
- Using arrays and Scalars
- Indexing Arrays
- Array Transposition
- Universal Array Function
- Array Processing
- Array Input and Output

Pandas:

- What is pandas ?
- Where it is used?
- Series in pandas
- Index objects
- Re index
- Drop Entry
- Selecting Entries
- Data Alignment
- Rank and Sort
- Summary Statics
- Missing Data
- Index Hierarchy

Scipy:

- Introduction
- Basic Functions
- Special functions (scipy.special)
- Integration (scipy.integrate)
- Optimization (scipy.optimize)
- Interpolation (scipy.interpolate)
- Fourier Transforms (scipy.fftpack)
- Signal Processing (scipy.signal)
- Linear Algebra (scipy.linalg)
- Sparse Eigenvalue Problems with ARPACK
- Compressed Sparse Graph Routines (scipy.sparse.csgraph)
- Spatial data structures and algorithms (scipy.spatial)
- Statistics (scipy.stats)
- Multidimensional image processing (scipy.ndimage)
- File IO (scipy.io)

Matplotlib: Python For Data Visualization

- Welcome to the Data Visualization Section
- Introduction to Matplotlib
- Plotting: how to
- The Matplotlib API
- Dealing with the third party libraries.
- The life cycle of plot
- Customizing Matplotlib with style sheets and rcParams

The Database toolkit for Python:

- Object Relational Tutorial
    - Mapper Configuration
    - Relationship Configuration

- Loading Objects
- SQL Sessions
- ORM Events and Internales

- SQLAlchemy Core
  - SQL Statements and Expressions API
  - SQLAlchemy schema metadata
  - Column and Data Types
  - Engine and Connection Use
- Dialects
  - Included Dialects
  - External Dialects

Python For Statistics:

Scikit: Scipy toolkit

Statistical learning: the setting and the estimator object in scikit-learn

- Datasets
- Estimators objects

Supervised learning: predicting an output variable from high-dimensional observations

- Nearest neighbor and the curse of dimensionality
- Linear model: from regression to sparsity
- Support vector machines (SVMs)

Model selection: choosing estimators and their parameters

- Score, and cross-validated scores
- Cross-validation generators
- Grid-search and cross-validated estimators

Unsupervised learning: seeking representations of the data

- Clustering: grouping observations together
- Decompositions: from a signal to components and loadings

Putting it all together

- Pipelining
- Face recognition with eigen faces
- Open problem: Stock Market Structure