DATA SCIENCE AND MACHINE LEARNING IN PYTHON

Nature of the Course: Theory + Practical
Total Hours per Day: 2 Hours
Course Duration: 8 Weeks

Course Summary

The DTC – Data Science and Machine Learning in Python course is targeted for beginners who want to learn how to think and write meaningful pieces of codes or read codes written by someone else in Python. This course teaches how to map literary description of a problem (requirement) to an application/library coded in Python. This is a core basic level course that is essential for anyone who has no prior programming experience but wishes to be a professional Python engineer in future.

Completion Criteria

After fulfilling all of the following criteria, the student will be deemed to have finished the module:

- Has attended 90% of all classes held.
- Has received an average grade of 80% on all assignments
- Has received an average of 60% in all assessments.
- The tutor believes the student has grasped all of the concepts and is ready to go on to the next module.

Required Textbooks

- Chris Albon, "Machine Learning with Python Cookbook", O'Reilly Media Inc. US.
- LazyProgrammer, "Deep Learning in Python", LazyProgrammer, US.
- Oliver Theobald, "Machine Learning for Absolute Beginners", Scatterplot Press.

Prerequisites

- Basic knowledge about programming, bits/bytes, procedures, classes, computer architecture, etc. If you just have theoretical knowledge that is perfectly okay but you should have strong convictions on what programming is, and what you hope to achieve from this class.
- Willing and eager to spend at least 10-20 hours (Varying from student-to-student) per week outside of the training class to self-study and practice.
- There is no prior educational level requirement for this course. Anyone from 10+2 student to someone who is doing their PHD in Genetic Engineering is welcome to take this course.
- If you are only interested in theory and have no interest/patience in spending at least 10 hours every week throughout the duration of the course, then this course might not be for you.

Course Details
Week I
Introduction To Programming Language
- Core Data Structures of Python
- Number
- String
- List
- Tuples
- Dictionary
- Set
- Advance Operation on Core Data-Structures
- Decision and Branching
- If, Else if, Else, Break, Continue
- Looping
- Functions
- Lambda Functions
- Map, Reduce, Filter
- Function Recursion
- Decorators

Week II
Python

- List and Dictionary Comprehension
- Exceptions and Exception Handling
- File Handling
- Object Oriented Programming (OOP)
- Introduction to Classes
- Inheritance, Encapsulation, Polymorphism, Abstraction
- Method Overloading
- Building Custom Packages and Modules

Week III
Basics To Data Science
- Introduction to Data Science
- Introduction to NumPy and Matplotlib
- Matrix Operations with NumPy
- Random Variable and Probability Distributions
- Probability
- Properties of Probability Distributions
- Mean, Median, Mode
- Variance, Skewness, Kurtosis
- Multivariate Normal Distribution
- Co-Variance, Correlation
- Introduction to Scikit-Learn
- Data Pre-Processing Techniques using Scikit-Learn
- Dimensionality Reduction as Data Pre-Processing
- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

Week IV
Machine Learning - I
- Introduction to Reinforcement Learning
- Q-Learning with Python
- Introduction to Clustering
- K-Means Clustering
- Agglomerative Clustering
- Introduction to Supervised Learning
- Naive Bayes Classification

Week V

Machine Learning – II

- Linear and Polynomial Regression
- K-Nearest Neighbors
- Decision Tree
- Balancing Bias vs Variance of ML Model
- Ensemble Learning
- Random Forest and Adaptive Boost
- Identifying Important Features of Data

Week VI

Deep Learning - I

- Introduction to Logistic Regression
- Computation Graph and Gradient Descent
- Introduction to Artificial Neuron (Perceptron)
- Multi-Layer Perceptron
- Introduction to Artificial Neural Networks
- Designing Artificial Neural Networks with Keras
- Gradient Descent Variants
- Classification and Regression using Neural Networks

Week VII

Deep Learning – II

- Introduction to Convolutional Neural Network (CNN)
- Object Classification with CNN
- Standard CNN Architectures
- Introduction to Object Detection
- The YOLO Algorithm

Week VIII

Natural Language Processing + Web Interface

- Introduction to NLTK
- Text Pre-Processing
- POS Tagging and Named-Entity Recognition
- Latent Semantic Analysis
- Introduction to Recurrent Neural Network
- Word2Vec Algorithm for Text Vectorization

- Natural Language Processing with LSTM
- Giving Web Interface to ML Application using Flask/Django

Labs

Lab assignments will focus on the practice and mastery of contents covered in the lectures; and introduce critical and fundamental problem-solving techniques to the students.

Learning Outcomes

- Understand concepts like lambdas and csv file manipulation.
- Describe common Python data science functionality and features.
- For cleaning and processing, learn how to use Data Frame structures.