AlgoProfessor ML/DATA Scientist Syllabus

Week 1, 2 & 3: Fundamentals of Python Programming, Mathematics, Probability and Statistics & optimization

Python for ML/Data Science:

Introduction ,Python, Anaconda and relevant packages installations, Why learn Python?

Keywords and Identifiers

Comments, Indentation, and Statements

Variables and Data types in Python

Standard Input and Output

Operators

Control flow: If...else , while loop ,for loop , break and continue

Lists ,Tuples ,Sets, Dictionary ,Strings

Types of Functions, Function Arguments

Recursive Functions, Lambda Functions

Mathematics :**Linear Algebra All Concepts w.r.to ML/Data Science**

Probability and Statistics: Introduction to Probability and Statistics.

Gaussian/Normal Distribution and its PDF(Probability Density Function).

CDF(Cumulative Density Function) of Gaussian/Normal Distribution

Standard normal variate (z) and standardization.

Sampling distribution & Central Limit Theorem

Discrete and Continuous Uniform distributions

Bernoulli and Binomial distribution

Hypothesis Testing methodology, Null-hypothesis, p-value

Optimization: All Concepts Related ML/ Data Science

Week 4, 5 &6: Python ML/Data Science Libraries:

NumPy, Pandas ,Matplotlib, Seaborn

NumPy :

Introduction to NumPy , NumPy Arrays ,

NumPy Indexing and Selection

NumPy Operations ,

Operations on NumPy Array

Pandas :

Introduction to Pandas, Series

DataFrames : Creating a DataFrame, Working with Columns, Working with Rows, Conditional Filtering

Missing Data ,GroupBy Operations

Combining DataFrames : Concatenation, Inner Merge, Left and Right Merge, Outer Merge

Pandas - Time Methods for Date and Time Data

Pandas Input and Output - CSV Files

Pandas Input and Output - HTML Tables

Pandas Input and Output - Excel Files

Pandas Input and Output - SQL Databases

Matplotlib :

Introduction to Matplotlib

Matplotlib Basics

Matplotlib - Understanding the Figure Object

Matplotlib - Figure Parameters

Matplotlib – SubPlots
Matplotlib Styling – Legends
Matplotlib Styling - Colors and Styles

Seaborn
Introduction to Seaborn
Scatterplots with Seaborn
Distribution Plots
Categorical Plots
Seaborn - Comparison Plots - Understanding the Plot Types
Seaborn Grid Plots
Seaborn - Matrix Plots

Week 7,8 &9: ML Supervised Algorithms With Examples

[Linear Regression, Logistic Regression, KNN, Support Vector Machines, Decision Tree, Random Forests, Boosting Methods]

Linear Regression & Project Based Model
Introduction to Linear Regression
Linear Regression - Understanding Ordinary Least Squares
Linear Regression - Cost Functions
Linear Regression - Gradient Descent
Python coding Simple Linear Regression
Linear Regression - Scikit-Learn Train Test Split
Linear Regression - Scikit-Learn Performance Evaluation – Regression
Linear Regression - Model Deployment and Coefficient Interpretation
Polynomial Regression - Creating Polynomial Features
Polynomial Regression - Training and Evaluation
Polynomial Regression - Choosing Degree of Polynomial
Polynomial Regression - Model Deployment
L1 and L2 Regularization - Elastic Net

Logistic Regression & Project Based Model

Introduction to Logistic Regression
Logistic Regression - Theory and Intuition
Logistic Regression with Scikit-Learn
Classification Metrics - Confusion Matrix and Accuracy
Classification Metrics - Precison, Recall, F1-Score
Classification Metrics - ROC Curves
Multi-Class Classification with Logistic Regression

KNN & Project Based Model
Introduction to KNN
KNN Classification - Theory and Intuition
KNN Coding with Python

Support Vector Machines

Introduction to Support Vector Machines

SVM - Theory and Intuition - Hyperplanes and Margins

SVM - Theory and Intuition - Kernel Intuition

SVM - Theory and Intuition - Kernel Trick and Mathematics

SVM with Scikit-Learn and Python – Classification

SVM with Scikit-Learn and Python – Regression

Decision Tree

Decision Tree - Understanding Gini Impurity

Constructing Decision Trees with Gini Impurity

Decision Trees with Python

Random Forests

Introduction to Random Forests Section

Random Forests - Key Hyperparameters

Random Forests - Bootstrapping and Out-of-Bag Error

Coding Classification with Random Forest Classifier

Coding Regression with Random Forest

Boosting Methods

Introduction to Boosting Section

AdaBoost Theory and Intuition, AdaBoost Coding

Gradient Boosting Theory

Gradient Boosting Coding

Week 10,11&12 : ML Unsupervised Algorithms With Examples
[K-Means Clustering, Hierarchical Clustering, DBSCAN, Principal Component Analysis, T-distributed stochastic neighborhood embedding]

K-Means Clustering
Introduction to K-Means Clustering

K-Means Clustering Theory

K-Means Clustering Coding

K-Means Color Quantization

Hierarchical Clustering
Introduction to Hierarchical Clustering

Hierarchical Clustering - Theory and Intuition

Hierarchical Clustering – Coding

DBSCAN

Introduction to DBSCAN

DBSCAN - Theory and Intuition

DBSCAN versus K-Means Clustering

DBSCAN - Hyperparameter Tuning Methods

DBSCAN – Python Code

Principal Component Analysis

Introduction to Principal Component Analysis
PCA Theory and Intuition
PCA - Manual Implementation in Python
PCA - SciKit-Learn

T-distributed stochastic neighborhood embedding (t-SNE)
What is t-SNE?

Neighborhood of a point, Embedding
t-SNE on MNIST

Week 13,14&15 : Reinforcement learning, Deep Learning Algorithms With Examples

Reinforcement Learning

The Markov decision process (MDP)
Types of Markov decision process
Reward vs Return
Discount factor
Bellman equations
Solving a Markov decision process
MDP in code
Reinforcement Learning Coding

Deep Learning
Deep Learning: Neural Networks.
Deep Learning: Deep Multi-layer perceptrons
Deep Learning: Tensorflow and Keras
Deep Learning: Convolutional Neural Nets
Deep Learning: Long Short-Term Memory (LSTMS)
Deep Learning generative Adversarial Networks(GANs)
Attention Models in Deep Learning
Deep Learning Real-World Case Studies

Dr Satyanarayana S