HiTech-Learner

Data Science and Machine Learning Using Python

Overview

Data Science and Machine learning is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence.

This Course will ensure that you become expert enough in Data Science and Machine Learning.The Course will cover 'Python concepts' and 'NLP - Basics' along with data science and machine learning. This course will also covers serving Machine Learning Model using RESTFull web service. You will be applying your learning in real time use cases. The course is designed keeping interviews level and questions in mind.

Course Details

I. Prerequisites

Candidate should have basic programming language knowledge. Though it's not mandatory. We will cover Python Programing language in the course.

Candidate should be from computer science background.

II. Duration

Course duration is approximately 5 Months for weekend classes and 3 months approximately for week days classes.

III. Training Mode - Online

Course will be conducted online via either Google Hangout or Go To Meeting. Organiser will schedule each training session and will send calendar invite to each candidate. Candidate can join the session by clicking on the meeting link given. Organiser will send reminder to join the meeting.

Each session will be recorded and recordings to be shared at the end of each session.

Candidates can raise their queries directly during the session and trainer would clarify each and every queries.

IV. Study Materials

Trainer would send class notes (explanations written on Jamboard) at the end of each session.

Trainer would also send course study material to every one either in the format of PPT or PDF. All practical exercises specially python files, jupyter notebook files would be shared to every candidate via google drive.

Candidate will get only View permission, no body except trainer can edit the study materials .

Validity of study materials is 6 months.

Our study materials are in depth and designed keeping interviews in mind.

Syllabus

➢ Python Programming

- Introduction to Python
- Basic Data Types
- Sequence Data Types
- Collection and Mapping Data Types
- Loops and Conditional Statements
- Functions , Iterators and Generators
- Python's inbuilt functions and modules
- OOPS concepts
- Handling Exceptions
- Packages and Modules
- Programing Exercises

➢ Numpy - Python Package

- Anaconda - Popular Scientific Distributions
- IPython Components
- Introduction to Jupyter Notebook
- Introduction to NumPy
- Creation of NumPy Array
- Array Shape Manipulation
- Basic Operations on NumPy Arrays
- Indexing, Slicing, Iterating Numpy Arrays
- Exercises on Numpy Array

➢ Panda - Data Analysis Tool

- Panda's Data Structures
- Accessing Data with Panda
- Knowing Data More
- Data Science and Machine Learning Using Python Nazrul Miya
- Python Programming 3 Sessions
- Numpy - Python Package 2 Sessions
- Panda - Data Analysis Tool 2 Sessions
- Data Frame
- Indexing Data Frame
- Data Cleaning

- Data Aggregation
- Data Merging
- Exercises with Real time DataSets

➢ Data Visualisation with Matplotlib & Seaborn

- introduction to Matplotlib
- Matplotlib Figure
- Line Plot and Scatter Plots
- Bar and Pie Plots
- Histograms and Box Plots
- Subplots
- Exercises

➢ Statistics and Probability

- Statistics and Data
- Statistics - Scenarios
- Probability Essentials
- Rules of Probability
- Random variables
- Expected Value and Variance
- Discrete Distribution
- Continuous Distribution
- Normal Distribution
- Naive Bayes Theorem
- Hypothesis Testing
- Chi-Squared Testing
- Multivariate Data Analysis
- Prior and Posterior Probability
- Markov Process
- Cumulative Density Function
- Kernel Density Estimation

➢ Linear Algebra

- Vector
- Matrices and Linear Transformations
- Matrix Multiplication
- Determinant of Matrix
- Matrix Dot and Cross Product

- Eigenvectors and Eigenvalues
- Singular Value Decompositions
- Exercises

➢ **Machine Learning Axioms**

- Supervised and Unsupervised Learning
- Features and Labels
- Regression and Classification
- Simple Linear Regression Model
- Multiple Linear Regression Model
- Logistic Regression
- Decision Tree
- Naive Bayes
- Gradient Descent
- Support Vector Machine
- Kernel Tricks
- K Nearest Neighbour (KNN)
- Clustering
- K-Means Clustering
- Neural networks
- Bagging and Boosting Algorithms
- Random Forest Algorithm
- Dataset Splitting - Training and Testing
- Cross validation
- Model Evaluation Techniques

All of the above algorithms will be shown in practical in real time use cases.

➢ **Natural Language Processing (NLP) Techniques**

➢ **Interview Question and Answer discussions**