Unati Global Connect Private Limited (UGCPL)

| Integrity | Sanctity | Commitment |

ONLINE TRAINING-CUM-INTERNSHIP PROGRAM POWERED BY UGCPL

Machine Learning using Python

Important information related to online Training-cum-Internship Program

- On registration confirmation, a common whatsapp group will be formed, where the students will be getting all necessary updates including the joining link (MS Teams / Google Meet) for online sessions.
- Two-hours online session on alternate–day–basis will be held from 3.00 pm to 5.00 pm (A minimum of 30 Hrs in one month). Timing schedule may change on request of majority of the participants
- 25 % of the allocated time will be for real time project work implementation.
- Project work will be in group of students (group will be consisting of a maximum of 8 students).
- Submission of complete project report by the participant is mandatory for the Internship Certification–One copy of the project report needs to be submitted at the parent Institution/Department and another copy will be required to be submitted to UGCPL.
- Software based project work will be free of cost and sufficient requirement for the Internship Certification. However, those students who wish to have their project work that need electronics/electrical/mechanical components, need to make their own arrangement for the same or alternatively bear the cost of the kit for delivery from UGCPL.
- Upon submission of project report in the concerned parent Institute/Department, the participants will be able to download their Internship Certificates within 10 working days from our website (www.ugcpl-india.com). The hard copy of the certificates will be submitted to the T&P Cell of the concerned University/Institute by UGCPL within said time period.
- After successful completion of Training–cum–Internship program, students may submit their copy of certificate along with their updated resume/CV online to UGCPL by e-mail for availing *lifetime free campus placement opportunities* offered by UGCPL.
- The participants will have to make their own arrangement of resources like Laptop or PC or smart phone and internet connectivity for attending the sessions through MS Teams/Google Meet platform.

Course Structure/Schedule of the Training-Cum-Internship on

MACHINE LEARNING USING PYTHON

Prerequisites

This training-cum-Internship Internship module is designed for learning from scratch. However, it is most suitable for students, developers, data analysts, and statisticians with basic knowledge of Computer Science and python programming.

Module 1

- The Machine Learning Landscape
- What Is Machine Learning?
- Why Use Machine Learning?
- Types of Machine Learning Systems
- Supervised/Unsupervised Learning
- Batch and Online Learning
- Instance-Based Versus Model-Based Learning
- Main Challenges of Machine Learning
- Insufficient Quantity of Training Data
- Non representative Training Data
- Poor-Quality Data
- Irrelevant Features
- Overfitting the Training Data
- Underfitting the Training Data
- Testing and Validating

Module 2

- NumPy Basics: Arrays and Vectorized Computation
- The NumPy ndarray: A Multidimensional Array Object
- Creating ndarrays
- Data Types for ndarrays
- Arithmetic with NumPy Arrays
- Basic Indexing and Slicing
- Boolean Indexing
- Fancy Indexing
- Transposing Arrays and Swapping Axes
- Universal Functions: Fast Element-Wise Array Functions
- Array-Oriented Programming with Arrays
- Expressing Conditional Logic as Array Operations
- Mathematical and Statistical Methods
- Methods for Boolean Arrays
- Sorting
- Unique and Other Set Logic
- File Input and Output with Arrays
- Linear Algebra
- Pseudorandom Number Generation
- Example: Random Walks

- Simulating Many Random Walks at Once

Module 3

- Getting Started with pandas Introduction to pandas Data Structures
- Series
- DataFrame
- Index Objects
- Essential Functionality
- Reindexing
- Dropping Entries from an Axix
- Indexing, Selection, and Filtering
- Integer Indexes
- Arithmetic and Data Alignment
- Function Application and Mapping
- Sorting and Ranking
- Axis Indexes with Duplicate Labels
- Summarizing and Computing Descriptive Statistics
- Correlation and Covariance
- Unique Values, Value Counts, and Membership

Module 4

- End-to-End Machine Learning Project
- Working with Real Data
- Look at the Big Picture
- Frame the Problem
- Select a Performance Measure
- Check the Assumptions
- Get the Data
- Create the Workspace
- Download the Data
- Take a Quick Look at the Data Structure
- Create a Test Set
- Discover and Visualize the Data to Gain Insights
- Visualizing Geographical Data
- Looking for Correlations
- Experimenting with Attribute Combinations
- Prepare the Data for Machine Learning Algorithms
- Data Cleaning
- Handling Text and Categorical Attributes
- Custom Transformers
- Feature Scaling
- Transformation Pipelines
- Select and Train a Model
- Training and Evaluating on the Training Set
- Better Evaluation Using Cross-Validation
- Fine-Tune Your Model
- Grid Search
- Randomized Search
- Ensemble Methods

- Analyze the Best Models and Their Errors
- Evaluate Your System on the Test Set
- Launch, Monitor, and Maintain Your System

Module 5

- Decision Tree
- Training and visualizing a decision tree
- Making predictions
- Estimating class probabilities
- The CART (Classification and Regression Tree) training algorithm
- Computational Complexity
- Gini Impurity or Entropy?
- Regularization hyper parameters
- Regression
- Instability

Module 6

- Dimensionality Reduction
- The Curse of Dimensionality
- Main Approaches for Dimensionality Reduction
- Projection
- Manifold Learning
- PCA
- Preserving the Variance
- Principal Components
- Projecting Down to d Dimensions
- Using Scikit-Learn
- Explained Variance Ratio
- Choosing the Right Number of Dimensions
- PCA for Compression
- Randomized PCA
- Incremental PCA
- Kernel PCA
- Selecting a Kernel and Tuning Hyperparameters
- LLE
- Other Dimensionality Reduction Techniques

Module 7

- Unsupervised Learning Techniques
- Clustering
- K-Means
- Limits of K-Means
- Using Clustering for Image Segmentation
- Using Clustering for Preprocessing
- Using Clustering for Semi-Supervised Learning
- DBSCAN
- Other Clustering Algorithms
- Gaussian Mixtures

- Anomaly Detection Using Gaussian Mixtures
- Selecting the Number of Clusters
- Bayesian Gaussian Mixture Models
- Other Algorithms for Anomaly and Novelty Detection

Project Work followed by project documentation