Industrial Summer Training

Name: Understanding the Machine Learning: An Industry Ready Approach
Duration: 7 Weeks (6 Weeks for Training and 1 week for Project)
Proposed Date: 21 June 2021 to 8 August 2021

This hands-on training enables participants to get a definitive understanding of Machine Learning Analytics Processing. It would incorporate basic building blocks of machine learning algorithms (Classification, Prediction, Segmentation, and Association) in deriving meaningful insights from the data. To apply the concepts learned on real world datasets through hands-on sessions using Python.

Session Duration: 3 hours/Day.
Course Fee: Rs. 4000/-

Beneficiaries: 2nd Year Students Seeking for Summer Industrial Training from: BE (CSE), BE (ECE), BE (EEE), BCA.

Registration Process: All the interested students have to register through the Google form Given below:

https://docs.google.com/forms/d/e/1FAIpQLSfIEdsUetUTvDYCQ5EqHrNjy7Kl9B-_buYHygKDUwFq7ywokQ/viewform?usp=sf_link

After registration you will be asked to pay the fee through CUIMS.

Last day of registration: 14 June 2021

Trainer: Dr. Piyush Samant and Mr. Gaurav Srivastava (AIT–CSE)

Learning Objectives:

- Understand the Basics of Python Programming.
- Understand Data and its type.
- Exploratory Data Analysis on Real time data.
- Statistical Analysis on Data.
- Knowledge on various machine learning Algorithms (MLA), tools and techniques to derive data driven insights
- Build and understand predictive (regression models) models.
- Build and understand classification models.
- Compare and Interpret Complex Models.
- Explore pattern through Clustering/Segmentation.
- An insight to Deep learning improvements.
- Apply the techniques learned through Python software

Data Analytics Tool Used: python. The course would be mainly taught using Python, industry preferred analytical tool most widely used by data scientists worldwide.

Target Audience: This is a beginner level course aimed at executives who are in the analytics domain(or) want to make a career shift to analytics.

Prerequisite: Basic knowledge of statistics. Knowledge about multivariate techniques such as factor analysis & multiple regressions would be an added advantage. The workshop would be taught using Python language; however, no working knowledge of Python is required.

Pedagogy: Pedagogy will be a mix of lecture and hands-on sessions using the appropriate analytics software/tools. For every concept that would be delivered, pedagogy would involve a right mix of theory sessions and hands on sessions using desktop/laptop. Appropriate real world datasets would be used for hands-on data analysis sessions.

Assessment Model: A student has to clear total 3 assessments to successfully complete the course (2 Internals and 1 Final)

Assessment 1	Assessment 2	Assessment 3
After successfully completion of the 1st and 2nd week	After successfully completion of the 3rd and 4th week	Final Assessment after successfully completing the course and submitting the project

Course Content in Details:

Module 1

1. Learning Paths
2. Machine Learning and its type.
3. ML vs. DL vs. AI. What's the Difference?
4. Introduction to Machine Learning Programming Environments.
5. Installation of Python Programming Environments.
6. Introduction to Python programming.
7. Python Functions.
8. Python OPPS Concepts.
9. Python Packages. (Pandas, Numpy, Seaborn)

Module 2

1. Exploratory Data Analysis (EDA): Univariate and Multivariate.
2. Visualization in EDA.
3. IQR and Z-Scores in EDA.
4. Probability foundation in EDA.
5. Hypothesis Testing.

Module 3

1. Linear Regression:
    a. Fundamentals of Linear Regression.
    b. Mathematics of Linear Regression.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.
2. Logistic Regression:

a. Fundamentals of Logistic Regression.
b. Mathematics of Logistic Regression.
c. Use Case Solving using python programming.
d. Performance Analysis of Model.

Module 4

1. Decision Tree:
    a. Fundamentals of Decision Tree.
    b. Mathematics of Decision Tree. (ID3, Entropy, IG, GI)
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.

2. Naïve Bayes Classifier:
    a. Fundamentals of Naïve Bayes Classifier.
    b. Mathematics of Naïve Bayes Classifier.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.

3. Support Vector Machine (SVM)
    a. Fundamentals of SVM.
    b. Mathematics of SVM.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.
4. KNN
    a. Fundamentals of KNN.
    b. Mathematics of KNN.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.

Module 5

1. Ensemble Techniques
    a. Fundamentals of Ensemble Techniques.
    b. Types of Naïve Bayes Classifier.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.

2. Group Based Clustering
    a. Fundamentals of Group Based Clustering Algorithms.
    b. Mathematics of Group Based Clustering Algorithms.
    c. Use Case Solving using python programming.
    d. Performance Analysis of Model.

3. Hierarchical Clustering
    a. Agglomerative
    b. Divisive
    c. Single Linkage, Complete Linkage, Ward's Method
    d. Fundamentals of Hierarchical Clustering.
    e. Mathematics of Hierarchical Clustering.

f. Use Case Solving using python programming.
g. Performance Analysis of Model.

Module 6

a. Upper Confidence Bound Fundamentals of Reinforcement Learning.
b. Application of Reinforcement Learning.
c. Use Case Solving using python programming.
d. Performance Analysis of Model.
e. AWS Deep Racer

Module 7

Final Project

Different Projects to individual students based on the real time industrial problems will be given based on their areas of interest.