DATA ANALYTICS & MACHINE LEARNING USING PYTHON

Nine-day Certificate Course by IIT Bombay

LEAD INSTRUCTOR

Prof. Asim Tewari

Professor (HAG), Center for Machine Intelligence and Data Science (C-MInDS) and Mechanical Engineering, IIT Bombay, Powai, Mumbai 400074

Table of Contents

INTRODUCTION

Why you should attend this program?

We have conducted 25 sessions till date and successfully trained more than 800 industry leaders and executives.

Every organization has data, but what makes the difference is how the data is put into action, be it to boost sales, reduce warranty costs, or increase the reliability of a design. Data analytics is the practice of deriving value out of data to drive decisions and actions. Data analytic techniques extract insights, patterns, and relationships in data by learning the structure of data. Common data analytic tasks include predictive modeling, data clustering, association rule mining, pattern mining, etc.

Often interpreting the analyzed data is more important than merely performing data analysis. Thus, data analytics requires a good knowledge of the various analysis techniques, and their underlying methodology, along with a basic understanding of how and when to apply these techniques.

This course will introduce and build a conceptual foundation of various popular data analytic techniques and will also provide training on performing each of these analytic tasks using Python programming through computer exercises and hands-on sessions. The hands-on exercise would be on real data sets. This course helps the participants to grasp the fundamentals of data analytics while also introducing them to the latest advancements in the area through theory, exercises, and case studies.

Who should attend?

Prerequisite

The program is exclusively designed for Junior to Middle-Level Executives. Intermediate knowledge of computer programming and mathematics is required to understand the course material.

PROGRAM OVERVIEW

What you will learn?

- Basic concepts and applications of various data analytic techniques for regression and classification
- Perform various data analytic tasks using Python programming
- Employ appropriate analytic techniques based on the problem definition at hand and also properly interpret the results of the analysis.
- Complete process of execution of a Data Analytics project from start to end.

Program Delivery

The course will be offered through online mode over six days. This will be followed by three days of offline training to conduct a Capstone project.

BROAD TOPICS

Key Areas Covered

Python

- Python overview
- Python tools
- Basic programming and data structures with python
- Data Handling tools in python

Data Science

- Problem framing
  - Understanding business problems and framing the problem in terms of data science and ML
- Data Science Life Cycle
  - Data Acquisition
  - EDA
  - Data Wrangling
    - Feature Engineering
  - Model Development

Machine Learning

- What is machine learning? And Types of Learning
  - Supervised
  - Unsupervised
  - Semi-supervised
  - Reinforcement
- Basic Probability and Statistics
  - Common probability distributions
  - The Law of Large Numbers
  - Tail bounds, Markov, Chebyshev, Chernoff, and Gaussian Annulus
- Linear Algebra
- Classification vs. Regression
- Parameters vs. Hyperparameters

BROAD TOPICS

Key Areas Covered

- Supervised Learning
  - Fundamental Algorithms
    - Linear Regression
    - Logistic Regression
    - KNN
    - Decision Trees
    - SVMs
- Regularized Models
  - Ridge Regression
  - Lasso Regression
- Ensemble Learning
  - Bagging vs Boosting
  - Gradient Boosting
  - Random Forest
  - Boosting Algorithms
    - AdaBoost
    - XGBoost
    - LGBM
    - Catboost
- Un-Supervised Learning
  - Clustering
    - KMeans
    - DBSCAN, HDBSCAN, BIRCH, and Affinity propagation
    - Semi-Supervised Learning
- Introduction to Deep Learning and Neural Networks

Best Practices

- Bias-Variance Trade-off
  - Bias, Variance
  - Overfitting vs. Underfitting
- Feature Selection
- Handling Imbalanced Dataset
- Curse of Dimensionality/Dimensionality Reduction
  - PCA, SVD, ICA
- Hyperparameter Tuning
  - Cross Validation techniques
- Ensemble Learning/Combining Models
  - Blending
  - Voting
- Stacking/Multi-Stage Models

BROAD TOPICS

Key Areas Covered

- Model Assessment
- Various metrics to monitor
- Which to use when?
- Model Explainability
- Experiments Tracking

Model Deployment

- Monitoring ML Models
- Different Types of Data Drifts:
  - Covariate Shift
  - Label/Target Distribution Shift
  - Concept Drift
- Re-trainable Pipeline
- Introduction to REST APIs
- Building REST APIs in Python
- Types of Deployment Strategies:
  - Batch Inference vs. Real-Time Inference
- ML Model Serialization and Deployment
- Docker for ML Model Deployment

Capstone project

- Define project objective
  - Define the scope of work
  - Identify deliverables
- Perform Literature review
- Understand data requirements
  - Data type
  - Data availability (internal, open source, etc.)
  - Data size and independence
- Data wrangling exploration
  - Data gathering (from multiple sources)
  - Formatting and blending
  - Cleanup and missing data
  - Data visualization and exploration
- AI Model development
  - Identify candidate AI models
  - Perform feature engineering and subset selection
  - Separate data into training and validation sets
  - Perform model training (if needed bootstrapping)

BROAD TOPICS

Key Areas Covered

- Model validation
  - Evaluate model results
  - MSE of training and testing
  - Perform hyperparameter tuning
- Model deployment with HMI
  - Develop HMI
  - Deploy model with HMI front end
  - Create basic test cases and user-help document
- Project report and presentation
  - Document the project process
  - Evaluate if the project objectives have been achieved
  - Provide suggestions for future improvements

About the Lead Instructor

LEAD INSTRUCTOR

Prof. Asim Tewari

Professor, Center for Machine Intelligence and Data Science (C-MInDS) and Mechanical Engineering, IIT Bombay, Powai, Mumbai 400074

Asim Tewari is a Professor(HAG) in the Department of Mechanical Engineering and a faculty member of the Center for Machine Intelligence and Data Science (C-MInDS) at the Indian Institute of Technology, Bombay, Mumbai. He graduated with a Bachelor's degree in Technology (B.Tech.) from IIT Kanpur and an M.S. and Ph.D. from Georgia Institute of Technology, Atlanta. With over twenty-five years of work experience in both corporate R&D and national research laboratories, he has made significant contributions to the field in Applied Artificial Intelligence in engineering systems and processes. He has published over 100 papers in international journals and conferences, as well as secured eleven international and ten Indian patents. Under his guidance, eighteen Ph.D. students have graduated, along with more than fifty M.S. students. Over the last decade, he has secured research funding of over $15 million.

At IIT Bombay, he has been instrumental in establishing several state-of-the-art facilities, including a Research Facility for Technical Textiles, a Medical Device Innovation Centre (BETiC), an Advanced Machining Excellence Cell, a Fiber Composite Research Laboratory, 4D X-ray Microscopy Laboratory, an experimental lab for thermo-mechanical simulation, and a Nano-characterization Texture Laboratory.

For the last decade, he has been working on the application of artificial intelligence in various domains. He established a Cyber-Physical Systems and Data Analytics research group at IIT Bombay, which focuses on smart manufacturing, machine learning, data analytics, and IoT for various sectors, including manufacturing, transportation, and defense. The group also works on video comprehension based on deep learning for inference and surveillance. Notably, the group has developed end-to-end solutions for digital factories using indigenous industrial IoT devices, cloud networking, and artificial intelligence backends. They have several international and Indian industrial projects to their credit and offer executive training on AI and digital transformation to industry.

FOR FURTHER DETAILS

Contact Us

iitb.ac.in

asim.tewari@iitb.ac.in

+91-22-2576-7521