



ASTHRA

TECHNICAL INSTITUTE

LIVE CLASSES

100% Placement Assistance

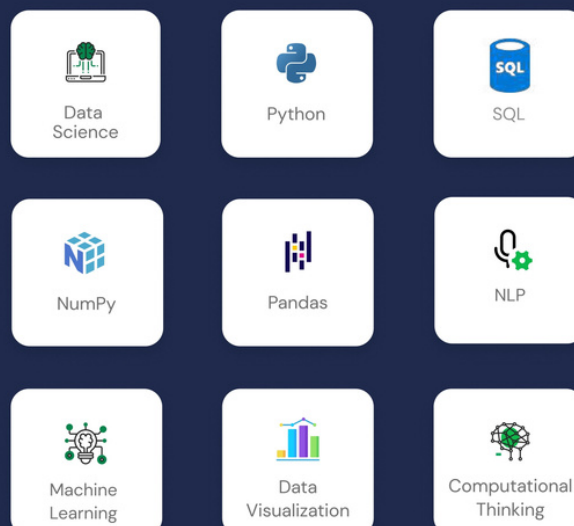
Data Science



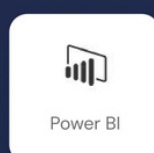
Top skills you'll learn!

- **Best Statistical programming language skills with Python.**
- **Excellent database querying skills.**
- **Good understanding of Analytical tools & Statistics**
- **Conceptual clarity towards Predictive performance & algorithm optimization.**
- **Master data visualization & communication skills.**

Technologies covered



Self-paced Courses – Add on



& more....

Module 0: Pre- Bootcamp

- This week you will go through the basics of competitive programming and data science as well as well we will come across the current industry trends in Data Science
- Introduction to Programming
- What is Data Science and what is Machine Learning?
- Competitive programming with Python
- (Codekata)
- Experts talk on industry trends
- Main Bootcamp flow

Program Curriculum

Module 1 : Python

We will go through the basics of python with all essential beginner friendly concepts of python programming like datatypes, loops, data structures and functions, followed by assessments and assignments

- Why python ?
- Python IDE
- Hello World Program
- Variables & Names
- String Basics
- List
- Tuple
- Dictionaries
- Conditional Statements
- For and While Loop , TRY AND EXCEPT
- Numbers and Math Functions
- Common Errors in Python

Module 2 : Python Advanced

Since we have essential basics of python we will see some advanced concepts like comprehensions, file handling, regular expressions, object oriented programming, pickling and many more essential concepts.

- Functions as ArgumentO
- List ComprehensioL
- File HandlinG
- Debugging in PythoL
- Class and ObjectO
- Lambda, Filters and MaU
- Regular ExpressionO
- Python PIA
- Read Excel Data in PythoL
- IteratorO
- PicklinG
- Python JSON

Module 3 : **Algorithmic thinking with Python**

We will explore the need of algorithmic thinking and the necessity of efficient coding, we will drive through data structures and algorithms along with memory management techniques

- **Introduction to algorithmic Thinking**
- **Algorithm Efficiency and time complexity**
- **Example algorithms - binary search, Euclid's algorithm**
- **Data structures - stack, heap, and binary trees**
- **Memory Management/Technologies**
- **Best Practices – Keeping it simple, dry code, naming Conventions, Comments, and docs.**
- **Assessment**

Module 4 : Data handling in Python - Pandas & MongoDB

Since we need to handle huge amounts of data, we will be implementing data handling techniques with Pandas library. And we will explore the different miscellaneous functions of Pandas library in detail.

- Introduction to Panda
- Series Data Structure - Querying and Indexing
- DataFrame Data Structure - Querying, Indexing, and loading
- Merging data frames
- Group by operation
- Pivot table
- Date/Time functionality
- Example: Manipulating DataFrame

Module 5 : SQL

we will dive into the SQL-based databases. We will learn the basics of SQL queries, schemas, and normalization.

- Joins(Inner,Left,Right,Full Join,Equi Join,Non-Equi Join,Self Join)
- Mathematical functions (SQRT,PI,SQUARE,ROUND,CEILING)
- Conversion functions(changing the data types)
- General functions(COALESCE,NVL,NULLIF)
- Conditional expressions (if,case,GO TO,NULL)
- Date and time function
- Numeric function
- String Function
- Subqueries
- Rank and Window Function
- Integrating Python with SQL

Module 6

We will Continue into the SQL-based databases.

We will learn the SQL Advanced queries, Join, Date and Time Functions and SubQueries.

- **Joins(Inner,Left,Right,Full Join,Equi Join, Non-Equi Join,Self Join)**
- **Mathematical functionV (SQRT,PI,SQUARE,ROUND,CEILING)**
- **Conversion functions(changing the data types)**
- **General functions(COALESCE,NVL,NULLIF)**
- **Conditional e;pressions (if,case,GO TO,NULL)**
- **Date and time functions**
- **Numeric functions**
- **String Functions**
- **Subqueries**
- **Rank and Window Functions**
- **Integrating Python with SQL**

Module 7 : Probability and Statistics with Numpy

We will go through Probability and Statistics whereas they are key to understanding, process and interpret the vast amount of data. We will deal with the basics of probability and statistics like Probability theory , Bayes theorem, distributions etc and their importance . Besides that we will do hands on with Numpy upon those concepts

- Why counting and probability theory
- Basics of sample and event space
- Axioms of probability
- Total Probability theorem and Bayes Theorem
- Random variables, PMF and CDF
- Discrete Distributions - Bernoulli, Binomial
- and Geometric
- Expectation and its properties
- Variance and its properties
- Continuous Distributions exponential and normal
- Sampling from continuous distributions
- Assessment

Module 8 : Probability and Statistics with Numpy - Continued

We will continue with statistics and probability and we will deal with descriptive and inferential statistics along with Hypothesis testing and lot of other relevant statistics methods

- Inferential statistics - sample vs population
- CLT and its proof
- Chi-squared distribution and its properties
- Point and Interval Estimators
- Estimation technique - MLE
- Interval Estimator of μ with unknown
- Examples of estimators
- Hypothesis testing - I
- Hypothesis testing - II
- Hypothesis testing - III
- Assessment

Module 9 : Data Visualisation in Python (Matplotlib/ Seaborn/ Plotly)

Data Visualization is used to understand data in visual context so that the patterns , trends and correlations in the data can be understood. We will do a lot of visualization with libraries like Seaborn, Matplotlib etc inturn that leads to effective story telling.

- Read Complex JSON files
- Styling Tabulation
- Distribution of Data - Histogram
- Box Plot
- Data Visualization - Recap
- Pie Chart
- Donut Chart
- Stacked Bar Plot
- Relative Stacked Bar Plot
- Stacked Qrea Plot
- Scatter Plots
- Bar Plot
- Assessment

Module 10 : **Data Engineering with Python**

It is always needed to analyze the data and preprocess it, since the real world data is not always industry ready, so in this week we will be dealing with a lot of data cleaning and Exploratory data Analysis techniques which is a very crucial stage for any data science project

- Handling missing data
- Techniques to impute missing values
- Encoding the data
- Outlier detection and correction
- Meaningful data transformation
- Assessment

Module 11 : Exploratory Data Analysis with Python

Real world data is always messy and it's very important to understand the statistical nature of data. Exploratory Data Analysis (EDA) is a critical step in the data analysis process, involving the preliminary examination of data to understand its characteristics, uncover patterns, and identify potential insights.

- **Descriptive Statistics:** Measures of central tendency (mean, median, mode); Measures of dispersion (range, variance, standard deviation); Skewness and kurtosis
- **Univariate Analysis:** Histograms, frequency distributions, and kernel density plots; Box plots and violin plots; Probability density functions (PDFs) and cumulative density functions (CDFs)
- **Bivariate Analysis:** Scatter plots and correlation analysis; Covariance and correlation coefficients; Pair plots and heatmaps
- **Multivariate Analysis:** PCA, Multivariate Scatter Plot, MANOVA
- **Real World Case Study**

Module 12 : Machine Learning with Sklearn

We are going to explore the need of machine learning and its types, Algorithms when to use and how to use essential mathematical intuition along with Evaluation metrics. We will see in detail about regression algorithms

- Introduction to machine learning
- Expert systems and 6 Jars
- Supervised Learning - Regression and Classification
- Evaluation metrics and measuring accuracy
- Introduction to regression
- Interpreting models
- Feature selection
- Regularization - Ridge and Lasso
- Assessment

Module 13 : Machine Learning with Sklearn - Continued

In continuation to the ML algorithms we are going to see in detail about different classification algorithms along with mathematical intuition and evaluation metrics

- Introduction to classification
- Evaluation metrics - TP, FP, and AUC
- Classification using logistic regression
- Classification using KNN
- Assessment

Module 14 : Machine Learning with Sklearn - Continued

We are going to explore classification algorithms like tree based algorithms in detail like how to interpret trees, pruning and ensemble methods like bagging and boosting etc

- Introduction to decision trees
- Building, pruning, and interpreting trees
- Ensemble techniques - Bagging and boosting
- Random forests
- Boosted trees - Gradient boosting
- Assessment

Module 15 : Machine Learning with Sklearn - Continued

After dealing with a lot of Supervised machine learning algorithms we will compare and get to know when to use what, Besides that we will deal with the do's and don'ts while training an ML model.

- Comparison of supervised techniques - when to use what?
- Comparison of supervised techniques - when to use what?
- Handling imbalanced data
- Undersampling
- Oversampling
- Other methods - ROSE, SMOTE, etc.
- Assessment

Module 16 : Machine Learning with Sklearn - Continued

Now we will explore Unsupervised learning algorithms, why unsupervised ?, when to use it and as well as the essential mathematical intuition

- Introduction to unsupervised learning
- Market Basket Analysis
- K means algorithm
- Assessment

Module 17 : **Deep learning**

As we move on to more complex problems, such as object recognition and text analysis, our data becomes extremely high dimensional, and the relationship becomes nonlinear. To accommodate this complexity, we move on to building more complex models that resemble our brain.

- **Fundamentals of Neural Networks: Limitations of ML; The Neuron; Linear perceptron as neurons**
- **Feed Forward Neural Networks: Linear Neurons and limitations; Sigmoid, Tanh and ReLU; Softmax**
- **Learning-I: Gradient Descent; Delta rule and learning rates; Gradient descent with sigmoidal Neurons**
- **Learning-II: Backpropagation; Stochastic and minibatch; Test set, validation set, and overfitting**
- **Preventing overfitting**

Module 18 : Deep learning with PyTorch

Now that we have a better theoretical understanding of deep learning models, we will spend this module implementing some of these algorithms in PyTorch

- PyTorch Basics: Installation and setup of PyTorch@ Tensors and operations in PyTorch
- Training Fundamentals: Autograd@ Backpropagation@ Gradient Descent@ Training Pipeline
- Regression with PyTorch: Linear Regression@ Logistic Regression
- Dataset in PyTorch: Dataset and Dataloader@ Dataset Transforms
- Training Pipeline: Softmax and Crossentropy? Z Activation Functions

Module 19 : Deep Learning with PyTorch continued

Now that we have the basic understanding of PyTorch, we will now dive into discussing the implementation details of a few state-of-the-art deep learning architectures in PyTorch

- Feed Forward Net: Creating basic Neural net@ Load Data and train neural net@ Evaluation on test set
- CNN: Introduction@ Image Filter/Image kernel@ Convolution layer and RGB@ Pooling Layer
- Transfer Learning
- Tensorboard
- Save and Load Models

Module 20 : **Natural Language Processing**

We are going to explore Natural Language Processing (NLP). Given the fact that we have a decent understanding of Machine Learning and Deep Learning, we can now explore the powerful ways to handle the NLP usecases

- Language Understanding: RNNs architectureD RNNs and language ModelsD Generation with RNNs
- Adding more memory LSTM architecture
- Encoder Decoder Model with RNN
- Self Attention Networks: Transformers
- Hands on Huggingface: Understanding API integration
- Using Language Models for various tasks: sentiPent analysisD Question AnsweringD NER summerization

Module 21 : Computer Vision

Having a basic understanding of NLP use cases, now we will dive into the Computer Vision Fundamentals. We will discuss state-of-the-art CV problems and their solutions with deep learning.

- Convolution Architecture: Filters, stacking Multiple Feature Maps, PyTorch Implementation
- Pooling Layers: Pytorch Implementation
- CNN Architectures overview: LeNet-5, AlexNet, GoogLeNet, VGGNet, ResNet, Xception, SENet
- Implementing a ResNet-34 CNN using PyTorch
- Using pretrained Models with PyTorch
- Object Detection: Fully Convolutional Networks, YOLO
- Semantic Segmentation

Module 22 : Model Deployment in AWS Cloud Platform

Having a good understanding of ML, DL and various use cases, we will now discuss the platforms through which we can securely deploy these powerful models on production level. More specifically; we will discuss the fundamentals of AWS services and how to use them efficiently.

- Introduction to AWS
- e Cloud Services (EC2, Lambda, S3, RDS etc)
- Hands-on in EC2 instance
- Hands-on in Database in AWS
- Hands-on in S3 storage
- Deploying ML Model as Application in AWS

Module 23 : Putting it together - Solving DS problems

This whole week we are going to work on industry projects which are currently in demand in the guidance of industry experts

- Case Study - I: Credit Card Fraud detection
- Case Study - II: Airline Customer segmentation
- Case Study - III: Product recommendation engine
- Case Study - IV: Chatbot with Huggingface

Module 24 : Mock Interviews

Eventually, it's time to attend the mock interviews which will be conducted by the industry experts like Data scientists, IIT professors and renowned HR's in order to mould you in every area possible



TECHNICAL INSTITUTE

Get Hired Once Course Got Completed

Be Part Of Upskill Ecosystem

Get Linkedin Etiquette Guide

Both Weekday & Weekend Classes are Available

100 % Placement Assistance



+91 7810077637



myasthra.com