Learnbay

Advanced Data Science & AI
Certification Program

300+ Hiring Partners

Hybrid Model for Project Sessions

175% Average Salary Hike

www.learnbay.co

+91 77956 87988

Learnvista Pvt. Ltd.

Context

About the program 01

Program highlights 02

Why learn data science? 03

Program details 04

Domain electives 05

Check out what our alumni have to say 06

Career service 07

Learning path 08

Program syllabus 09

Real-time projects and case-studies 10

Certification 11

About The Program

Our Advanced Data Science and AI program is designed for professionals seeking to advance their careers in the fast-changing field. It offers a holistic learning experience with practical applications and real-world case studies. The program covers cutting-edge technologies and ethical considerations in AI to prepare working professionals for success.

2cr
worth scholarships awarded

600+
professionals secured jobs after a career break

30k+
Trusted Learners

We exist to provide accessible, reasonable, and industry-relevant education that empowers India's workforce to grow and develop.

4.79/5 4.66/5 4.8/5

Great Place To Work. Certified

Thousands of student reviews on Switchup, Course Report, Google and more

Program Highlights

Industry-Relevant & Updated Syllabus

Learn new tools, techniques & trends. Gain app development experience.

360 Degree Knowledge Building

Develop practical skills through real-world projects and assignments.

One on one Dedicated Mentorship

Personalized learning experience from experienced industry professionals.

Multiple Career Opportunities

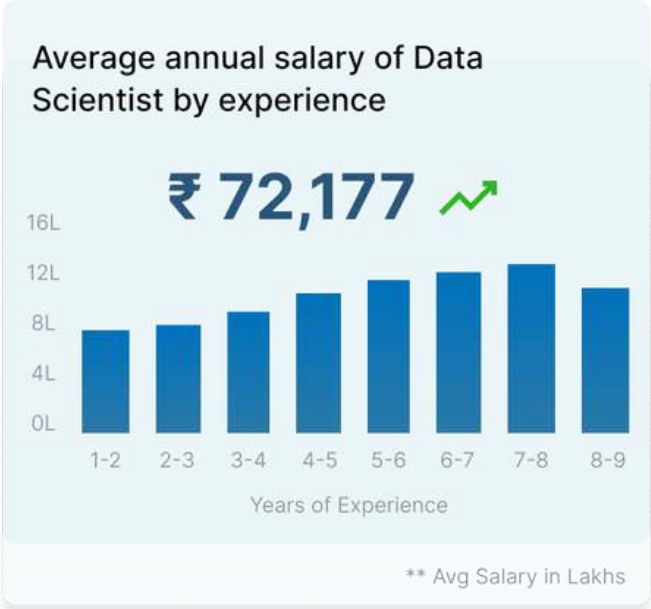
Advance in data science career by pursuing roles like data scientist, engineer, etc.

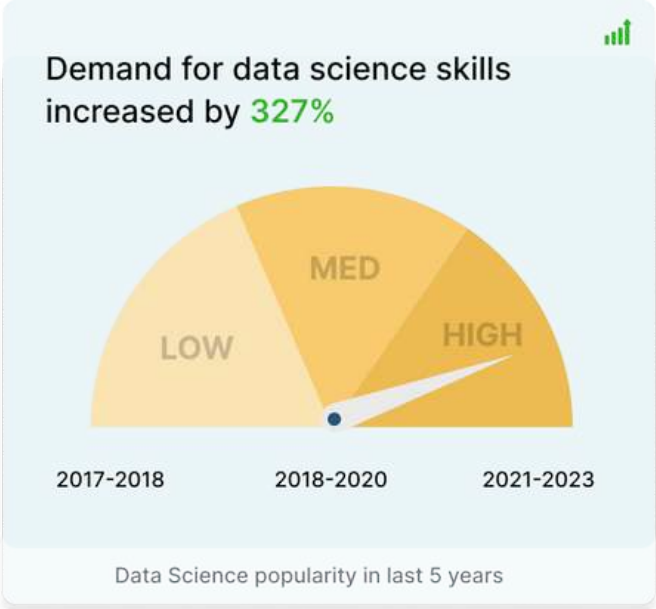
Why Learn Data Science?

60% rise in data science jobs

250% highest salary hike

300+ partner companies





Placement Report

30K+
Trusted Learners

9K+
Successfully Placed

50K+
Job Interviews Cracked

Book a free consultation with expert **Contacts Us**

Program Details

ELIGIBILITY

- **Working professional** having more than 6 months of experience in any domain (Technical/Non-Technical)

Qualification:

BE/B.Tech (from any branch), BBA/MBA, MCA/M.Tech, B.Com, B.Sc (in any branch)

Course duration : 250+ hours

Weekday Batch : 7 months	**Weekend Batch : 9 months**
Monday - Friday : 2 hrs/day	**Saturday - Sunday: 3.5 hrs/day**

About instructors:

Experienced software development instructors share valuable practical knowledge and effective solutions, preparing students for success in the industry.

TOTAL FEES

₹ 99,000/- + 18% GST

₹ 1,16,820/-

EASY EMI

₹ 9,735/month

Financing partners

zest ShopSe propelld

Domain Electives

Select any 1 domain and become a domain expert

BFSI

Master financial analysis for strategic development success

Healthcare

Use advanced tools & methods to gain a competitive edge

Manufacturing

Explore advanced tech for strategic skill development

Energy, Oil & Gas

Excel in data analysis, craft dynamic dashboard for insights

Supply chain, E-com & Retail

Master supply chain domain with data-driven insights and strategies

What Our Alumni's Say

Shravanthi A

Data Scientist

Ford

Learnbay has helped me a lot to learn data science applications in the e-commerce industry. The live class concept was really helpful in receiving proper DS training. Thanks to all my mentors and the placement team.

230%
Salary Hike

Ritesh Kumar

Associate Consultant

Capgemini

I knew nothing about data science before I joined Learnbay. But through a variety of instructors, I steadily developed my notion and received solid knowledge and conceptual training in data science with hike of 150%.

150%
Salary Hike

Saurabh Kumar

Data Scientist & Statistician

Teleperformance

When I joined Learnbay I did not have any knowledge apart from the very basics. I gradually build my concept via various trainers and get trained in data science with strong knowledge/concepts.

135%
Salary Hike

What Our Alumni's Say

Preksha Mishra

Lead Data Scientist

HCL

The course structure is excellent with emphasis on concept building and tools & software at the same time. The support team is excellent and supportive and quite agile to respond to doubts.

140%
Salary Hike

Mohd. Israr

Data Scientist

Thanks to the Learnbay data science course & excellent guidance, I was able to ace the TCS interview and secure a job with a 210% pay raise. The real-world time projects helped me develop my concepts as a data scientist.

210%
Salary Hike

Sweekrithi shetty

Learnbay stands out with expert-led courses that offer practical and up-to-date content. Hands-on learning and ample resources make it accessible for anyone seeking to upskill in AI, ML, or data science. Learnbay is an excellent choice to learn and grow in these rapidly-evolving fields.

Certificates

World's leading certifcations

IBM

IBM Course Certificate

Complete your training with the globally recognized certificate.

Microsoft

Microsoft Course Certificate

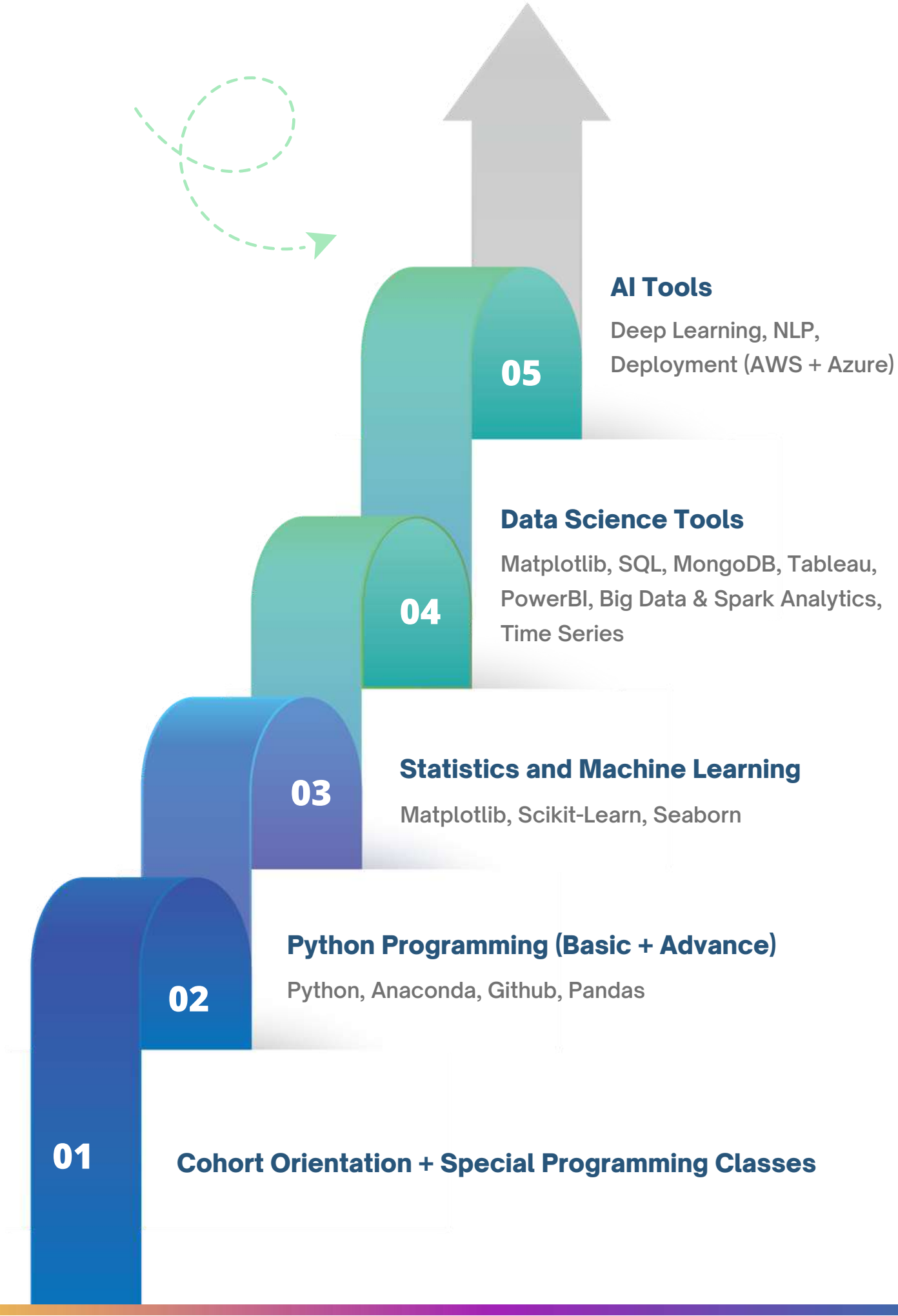
Achieve professional growth & increase earning potential with Microsoft certification

IBM

IBM Capstone Project Certificate

Highlight your skills & boost your project portfolio with capstone project certificate

Learning Path



Career Service

Get 1 year of Job and Placement support

Unleash your career potential with unlimited job access, interview support, and profile review.

Get 3 mock interviews with industry leaders

Master the art of data science & AI and stay ahead of the curve with mockups and industry insights

Resume build up session

Craft a powerful resume showcasing your expertise in data science and AI to stand out from the competition.

Get 5-8 interview calls

Receive 5-8 interview calls from a diverse pool of interested employers/recruiters.

Others Vs Learnbay

Benefits	Learnbay	Others
Guaranteed Interview Calls	✓	✗
Industry capstone project certificate from IBM	✓	✗
Domain specialized programs for professionals	✓	✗
100% live interactive sessions with industry experts	✓	✗
On demand video call with industry experts	✓	✗
Personalized Resume Review Session	✓	✗

Program Syllabus

Preparatory Session

Module 0 (08 hours)

Preparatory Session

- A brief introduction to tools related to data
- Learn about particular real-time projects and Capstone projects
- Data and its impact on career opportunities
- Fundamental relevance of projects using data
- Role of data in businesses
- Significance of data in decision-making
- Scope of data in research and development
- Utilizing data, to enhance industrial operations and management
- Data in performance evaluation
- Data in customer segmentation

Fundamentals of Statistics

- Mean, Median, Mode
- Standard Deviation, Average.
- Probability, permutations, and combinations
- Introduction to Linear Algebra

Fundamentals of programming

- Types of code editors in python
- Introduction to Anaconda & Jupyter notebook
- Flavors of python
- Introduction to Git, GitHub
- Python Fundamentals
- Source code vs Byte code vs Machine code
- Compiler & Interpreter
- Memory Management in Python

Tools covered

ANACONDA jupyter git GitHub

Program Syllabus

TERM 1

Python Programming

Module 1 (50 hours)

Programming Basics & Environment Setup

- Installing Anaconda, Anaconda Basics and Introduction
- Get familiar with version control, Git and GitHub.
- Basic Github Commands.
- Introduction to Jupyter Notebook environment. Basics Jupyter notebook Commands.
- Programming language basics

Strings, Decisions & Loop Control

- Working With Numbers, Booleans
- and Strings, String types and formatting, String operations
- Simple if Statement, if-else Statement
- if-elif Statement.
- Introduction to while Loops, for Loops, Using continue and break

Class Hands-on:

- 6 programs/coding exercise on string, loop and conditions in classroom

Functions And Modules

- Introduction To Functions
- Defining & Calling Functions
- Functions With Multiple Arguments

Python Programming Overview

- Python Overview
- Python 2.7 vs Python 3
- Writing your First Python Program
- Lines and Indentation, Python Identifiers
- Various Operators and Operators Precedence
- Getting input from User, Comments, Multi line Comments

Python Data Types

- List, Tuples, Dictionaries
- Python Lists, Tuples, Dictionaries Accessing Values, Basic Operations
- Indexing, Slicing, and Matrixes
- Built-in Functions & Methods
- Exercises on List, Tuples And Dictionary

Functions And Modules

- Anonymous Functions - Lambda
- Using Built-In Modules, User-Defined Modules, Module Namespaces,
- Iterators And Generators

Class Hands-on:

8+ Programs to be covered in class of functions, Lambda, modules, Generators and Packages.

Program Syllabus

TERM 1

Python Programming

Module 1 (50 hours)

File I/O An d Exceptional Handling and Regular Expression

- Opening and Closing Files
- open Function, file Object Attributes
- close() Method, Read, write, seek.
- Exception Handling, try-finally Clause
- Raising an Exceptions, User-Defined Exceptions
- Regular Expression- Search and Replace
- Regular Expression Modifiers
- Regular Expression Patterns

Class hands-on :

- 10+ Programs to be covered in class from File IO, Reg-ex and exception handling.

Data Analysis Using Pandas

- Pandas : Introduction to Pandas
- Importing data into Python
- Pandas Data Frames, Indexing Data Frames ,Basic Operations With Data frame, Renaming Columns, Subsetting and filtering a data frame.

Data Analysis Using Numpy

- Introduction to Numpy. Array Creation, Printing Arrays, Basic Operation - Indexing, Slicing and Iterating, Shape Manipulation - Changing shape, stacking and splitting of array
- Vector stacking, Broadcasting with Numpy, Numpy for Statistical Operation

Assignment 1 (Week 2):

10 Coding exercises on Python Basics - Variables, Operators, Strings, Loops, Control Statement

Assignment 2 (Week 3):

10 Python programs and practice set on List, Tuples, Dictionaries & Matrices operations

Assignment 3 (Week 4):

10 Coding exercises on Functions, Lambda, Input-Output, File and Regular Expression

Program Syllabus

TERM 1

Python Programming

Module - 1 (50 hours)

Data Visualization using Matplotlib

- **Matplotlib:** Introduction, plot(), Controlling Line Properties, Subplot with Functional Method, Multiple Plot, Working with Multiple Figures, Histograms

Data Visualization using Seaborn

- **Seaborn:** Intro to Seaborn And Visualizing statistical relationships , Import and Prepare data. Plotting with categorical data and Visualizing linear relationships.
- Seaborn Exercise

Case Study

3 Case Study on Numpy, Pandas, Matplotlib

1 Case Study on Pandas And Seaborn

Assessment Test in Python :

- 2 hour of Assesment Test in Python (Coding & Objective Questions)

Real time Use cases in Python to be Covered in Class with 5 assignments

NumPy pandas matplotlib seaborn

Program Syllabus

TERM 2

Statistics

Module - 1 (30 hours)

Fundamentals of Math and Probability

- Probability distributed function & cumulative distribution function. Conditional Probability, Baye's Theorem
- Problem solving for probability assignments
- Random Experiments, Mutually Exclusive Events, Joint Events, Dependent & Independent Events

Introduction to Statistics, Statistical Thinking

- Variable and its types
- Quantitative, Categorical, Discrete, Continuous,
- *all with examples

Five Point Summary and Box Plot

- Outliers, Causes of Outliers, How to treat Outliers, I-QR Method and Z-Score Method

Inferential Statistics

- Central Limit Theorem
- Point estimate and Interval estimate
- Creating confidence interval for population parameter

All about Population & Sample

- Population vs Sample, Sample Size
- Simple Random Sampling, Systematic Sampling, Cluster Sampling, Stratified Sampling, Convenience Sampling, Quota Sampling, Snowball Sampling and Judgement Sampling

Descriptive Statistics

- Measures of Central Tendency – Mean, Median and Mode
- Measures of Dispersion – Standard Deviation, Variance, Range, IQR (Inter-Quartile Range)
- Measure of Symmetricity/ Shape – Skewness and Kurtosis

Inferential Statistics

- Characteristics of Z-distribution and T-Distribution.
- Type of test and rejection region.
- Type of errors in Hypothesis Testing

Program Syllabus

TERM 2

Statistics

Module - 1 (30 hours)

Hypothesis Testing

- Type of test and Rejection Region
- Type o errors-Type 1 Errors, Type 2 Errors. P value method, Z score Method. The Chi-Square Test of Independence.
- Regression. Factorial Analysis of Variance. Pearson Correlation Coefficients in Depth. Statistical Significance
- Null and Alternative Hypothesis One-tailed and Two-tailed Tests, Critical Value, Rejection region, Inference based on Critical Value
- **Binomial Distribution:** Assumptions of Binomial Distribution, Normal Distribution, Properties of Normal Distribution, Z table, Empirical Rule of Normal Distribution & Central Limit Theorem and its Applications

Data Processing & Exploratory Data Analysis

- What is Data Wrangling
- Data Pre-processing and cleaning?
- How to Restructure the data?
- What is Data Integration and Transformation

Linear Algebra

- Dot Product, Projecting Point on Axis.
- Matrices in Python, Element Indexing, Square Matrix, Triangular Matrix, Diagonal Matrix, Identity Matrix, Addition of Matrices, Scalar Multiplication, Matrix Multiplication, Matrix Transpose, Determinant, Trace
- T-Test, Analysis of variance (ANOVA), and Analysis of Covariance (ANCOVA) Regression analysis in ANOVA

Class Hands-on:

- Problem solving for C.L.T Problem solving Hypothesis Testing Problem solving for T-test, Z-score test Case study and model run for ANOVA, ANCOVA

Program Syllabus

TERM 2

Statistics

Module - 1 (30 hours)

EDA

- Finding and Dealing with Missing Values.
- What are Outliers?
- Using Z-scores to Find Outliers.
- Bivariate Analysis, Scatter Plots and Heatmaps.
- Introduction to Multivariate Analysis

Note: Problem-Solving Techniques and Case Studies using Statistics will be covered in class from week 2

Statistics Assignments : Total 4 practice set and Assignments from Statistics

Program Syllabus

TERM 2

Machine Learning

Module - 2 (40 hours)

Machine Learning Introduction

- Definition, Examples, Importance of Machine Learning
- Definition of ML Elements: Algorithm, Model, Predictor Variable, Response Variable, Training - Test Split, Steps in Machine Learning,
- ML Models Type: Supervised Learning, Unsupervised Learning and Reinforcement Learning

Data Preprocessing

- Types of Missing values (MCAR, MAR, MNAR) , Methods to handle missing values
- Outliers, Methods to handle outliers: IQR Method, Z Method
- Feature Scaling: Definition , Methods: Absolute Maximum Scaling, Min-Max Scaler , Normalization, Standardization, Robust Scaling

Data Preprocessing

- Encoding the data: Definition, Methods: OneHot Encoding, Mean Encoding, Label Encoding, Target Guided Ordinal Encoding

Logistic Regression Model

- Definition. Why is it called the “Regression model”?
- Sigmoid Function, Transformation & Graph of Sigmoid Function

Evaluation Metrics for Classification model

- Confusion Matrix, Accuracy, Misclassification, TPR, FPR, TNR, Precision, Recall, F1 Score, ROC Curve, and AUC. Using Python library Sklearn to create the Logistic Regression Model and evaluate the model created

K Nearest Neighbours Model

- Definition, Steps in KNN Model, Types of Distance: Manhattan Distance, Euclidean Distance, ‘Lazy Learner Model’.
- Confusion Matrix of Multi Class Classification
- Using Python library Sklearn to create the K Nearest Neighbours Model and evaluate the model

Program Syllabus

TERM 2

Machine Learning

Module - 2 (40 hours)

Decision Tree Model

- Definition, Basic Terminologies, Tree Splitting Constraints, Splitting Algorithms:
- CART, C4.5, ID3, CHAID
- Splitting Methods:
- GINI, Entropy, Chi-Square, and Reduction in Variance
- Using Python library Sklearn to create the Decision Tree Model and evaluate the model created

Hyperparameter Tuning

- GridSearchCV, Variable Importance.
- Using Python library Sklearn to create the Random Forest Model and evaluate the model created.
- Use cases

Random Forest Model

- Ensemble Techniques: Bagging/bootstrapping & Boosting.
- Definition of Random Forest, OOB Score
- K-Fold Cross-Validation

Naive Baye's Model

- Definition, Advantages, Baye's Theorem Applicability, Disadvantages of Naive Baye's Model, Laplace's Correction, Types of Classifiers: Gaussian, Multinomial and Bernoulli
- Using Python library Sklearn to create the Naive Baye's Model and evaluate the model created

Case Study

- **Business Case Study for Kart Model**
- **Business Case Study for Random Forest**
- **Business Case Study for SVM**
- **To classify an email as spam or not spam using logistic Regression.**
- **Application of Linear Regression for Housing Price Prediction**

Program Syllabus

TERM 2

Machine Learning

Module - 2 (40 hours)

K Means and Hierarchical Clustering

- Definition of Clustering, Use cases of Clustering
- K Means Clustering Algorithm, Assumptions of K Means Clustering
- Sum of Squares Curve or Elbow Curve

Hierarchical Clustering

- Dendrogram, Agglomerative Clustering, Divisive Clustering, Comparison of K Means Clustering and Hierarchical Clustering
- Using Python library Sklearn to create and evaluate the clustering model

Principal Component Analysis(PCA)

- Definition, Curse of Dimensionality, Dimensionality Reduction Technique, When to use PCA,
- Use Cases
- Steps in PCA, EigenValues and EigenVectors, Scree Plot.
- Using Python library Sklearn to create Principal Components

Support Vector Machine(SVM)

- Model: Definition, Use Cases, Kernel Function, Aim of Support Vectors, Hyperplane, Gamma Value, Regularization Parameter
- Using Python library Sklearn to create and evaluate the SVM Model

Summary of all Machine Learning Models and Discussion about the Capstone Project

Note : All Machine Learning Algorithms are covered in depth with real time case studies for each algorithm. Once 60% of ML is completed, Capstone Project will be released for the batch.

Program Syllabus

TERM 2

CASE STUDY

Module - 2 (40 hours)

- **Recommendation Engine for e-commerce/retail chain**
- **Twitter data analysis using NLP**

SciPy

learn

Program Syllabus

TERM 3

SQL

Module - 1 (14 hours)

SQL and RDBMS

- RDBMS And SQL Operations.
- Single Table Queries - SELECT, WHERE,
- ORDER BY, Distinct, And, OR
- Multiple Table Queries: INNER, SELF,
- CROSS, and OUTER, Join, Left Join, Right
- Join, Full Join, Union

NoSQL, HBase & MongoDB

- NoSQL Databases
- Introduction to HBase
- HBase Architecture, HBase
- Components, Storage Model of HBase
- HBase vs RDBMS
- Introduction to Mongo DB, CRUD
- Advantages of MongoDB over RDBMS

Programming with SQL

- Mathematical Functions
- Variables
- Conditional Logic
- Loops
- Custom Functions
- Grouping and Ordering

Advance SQL

- Advance SQL Operations
- Data Aggregations and summarizing the data
- Ranking Functions: Top-N Analysis
- Advanced SQL Queries for Analytics

JSON Data & CRUD

- Basics and CRUD Operation
- Databases, Collection & Documents
- Shell & MongoDB drivers
- What is JSON Data
- Create, Read, Update, Delete
- Finding, Deleting, Updating, Inserting Elements
- Working with Arrays
- Understanding Schemas and Relations

Programming with SQL

- Partitioning
- Filtering Data
- Subqueries

Program Syllabus

TERM 3

SQL

Module - 1 (14 hours)

Assignments

- Working with multiple tables
- Practice Joins, Grouping and Subqueries
- Using GROUP BY and HAVING Clauses
- Practice Aggregation Queries

Program Syllabus

TERM 3

MongoDB

Module - 2 (14 hours)

Introduction to MongoDB

- What is MongoDB
- Characteristics and Features
- MongoDB Ecosystem
- Installation process
- Connecting to MongoDB database
- Introduction to NoSQL
- Introduction of MongoDB module
- What are Object Ids in MongoDB

Assignment

- Obtain the data in the format you want by formulating queries that are both effective and high-performing.

MongoDB (Advance)

- MongoDB Use cases
- MongoDB Structures
- MongoDB Shell vs MongoDB Server
- Data Formats in MongoDB
- MongoDB Aggregation Framework
- Aggregating Documents
- Working with MongoDB Compass & exploring data visually
- Understanding Create, Read, Update, Delete
- Schemas & Relations
- Document Structure
- Working with Numeric Data
- Working on Scheme Designing

Tools covered

mongoDB.

Program Syllabus

TERM 3

Tableau

Module - 3 (14 hours)

Introduction to Tableau

- Connecting to data source
- Creating dashboard pages
- How to create calculated columns
- Different charts

Dashboard and Stories

- Working in Views with Dashboards and Stories
- Working with Sheets
- Fitting Sheets
- Legends and Quick Filters
- Tiled and Floating Layouts, Floating Objects

Hands-on Assignments

- Connecting data source and data cleansing
- Working with various charts
- Deployment of Predictive model in visualization

Visual Analytics

- Getting Started With Visual Analytics
- Sorting and grouping
- Working with sets, set action
- Filters: Ways to filter, Interactive Filters
- Forecasting and Clustering

Tableau (Advance)

- Mapping
- Coordinate points
- Plotting Latitude and Longitude
- Custom Geocoding
- Polygon Maps
- WMS and Background Image

Tools covered

tableau

Program Syllabus

TERM 3

PowerBI

Module - 4 (14 hours)

Getting Started With Power BI

- Installing Power BI Desktop and Connecting to Data
- Overview of the Workflow in Power BI Desktop
- Introducing the Different Views of the Data Mode
- Query Editor Interface
- Working on Data Model

Programming with Power BI

- Working with Time Series
- Understanding aggregation and granularity
- Filters and Slicers in Power BI Maps
- Scatterplots and BI Reports
- Connecting Dataset with Power BI Creating a Customer Segmentation Dashboard Analyzing the Customer Segmentation Dashboard

Assignments

- Create Bar charts
- Create Pie charts
- Create Tree maps
- Create Donut Charts
- Create Waterfall Diagrams
- Creating Table Calculations for Gender

Tools covered

Power BI

Program Syllabus

TERM 3

Big Data & Sparks Analytics

Module - 5 (16 hours)

Introduction To Hadoop & Big Data

- Distributed Architecture - A Brief Overview. Understanding Big Data
- Introduction To Hadoop, Hadoop Architecture
- HDFS, Overview of MapReduce Framework
- Hadoop Master: Slave Architecture
- MapReduce Architecture
- Use cases of MapReduce

What is Spark

- Introduction to Spark RDD
- Introduction to Spark SQL and Data frames
- Using R-Spark for machine learning
- Hands-on:
- Installation and configuration of Spark
- Using R-Spark for machine learning programming

Hands-on

- Map reduce Use Case 1: Youtube data analysis
- Map reduce Use Case 2: Uber data analytics
- Spark RDD programming
- Spark SQL and Data frame programming

Tools covered

Spark hadoop SQL PySpark

Program Syllabus

TERM 3

Time Series

Module - 6 (14 hours)

Introduction to Time Series Forecasting

- Basics of Time Series Analysis and Forecasting
- Method Selection in Forecasting
- Moving Average (MA) Forecast Example
- Different Components of Time Series Data
- Log Based Differencing, Linear Regression for Detrending

Introduction to ARIMA Models

- ARIMA Model Calculations, Manual ARIMA Parameter Selection
- ARIMA with Explanatory Variables
- Understanding Multivariate Time Series and their Structure
- Checking for Stationarity and Differencing the MTS

CASE STUDY

- **Time series classification of smartphone data to predict user behavior**
- **Performing Time Series Analysis on Stock Prices**
- **Time series forecasting of sales data**

Note: All the assignments and case studies will be covered in-depth with real-time examples

Program Syllabus

TERM 4

Deep Learning Using Tensorflow

Module - 1 (20 hours)

Introduction to Deep Learning And TensorFlow

- Neural Network
- Understanding Neural Network Model
- Installing TensorFlow
- Simple Computation, Constants, and Variables
- Types of file formats in TensorFlow
- Creating A Graph – Graph Visualization
- Creating a Model – Logistic Regression
- Model Building using tensor flow

Understanding Neural Networks With TensorFlow

- Basic Neural Network
- Single Hidden Layer Model
- Multiple Hidden Layer Model
- Backpropagation – Learning Algorithm and visual representation
- Understand Backpropagation – Using Neural Network Example
- TensorBoard

TensorFlow Classification Examples

- Introduction to TensorFlow
- Installing TensorFlow
- Simple Computation, Contents
- and Variables
- Types of file formats in TensorFlow
- Creating A Graph - Graph Visualization
- Creating a Model - Logistic Regression Model Building
- TensorFlow Classification Examples

Convolutional Neural Network (CNN)

- Convolutional Layer Motivation
- Convolutional Layer Application
- The architecture of a CNN
- Pooling Layer Application
- Deep CNN
- Understanding and Visualizing a CNN

Project

- Building a CNN for Image Classification
- Project on backpropagation using Neural Networks with Tensor Flow

Program Syllabus

TERM 4

Natural Language Processing

Module - 2 (24 hours)

Natural Language Processing

- Text Analytics
- Introduction to NLP
- Use cases of NLP algorithms
- NLP Libraries
- Need for Textual Analytics
- Applications of NLP
- Word Frequency Algorithms for NLP Sentiment Analysis

Text Analysis

- Distance Algorithms used in Text Analytics
- String Similarity
- Cosine Similarity Mechanism -
- The similarity between two text documents
- Levenshtein distance - measuring the difference between two sequences

Important

- Applications of Levenshtein distance
- LCS(Longest Common Sequence)
- Problems and solutions, LCS Algorithms

KNN

- Information Retrieval Systems
- Information Retrieval - Precision, Recall,F- score TF-IDF
- KNN for document retrieval
- K-Means for document retrieval
- Clustering for document retrieval

Use cases on NLP

- Sentiment analysis for marketing
- Toxic comments classification
- Language identification
- Generating research papers titles
- Application to translate and summarize the news
- RESTful API for similarity check

Program Syllabus

TERM 4

Deployment AWS+Azure

Module - 3 (10 hours)

Introduction to AWS and Azure Machine Learning Services

- Overview of AWS SageMaker and Azure Machine Learning
- Key features and benefits of using these platforms
- Understanding different types of machine learning algorithms and use cases

Data Preparation and Feature Engineering

- Understanding the data requirements for machine learning models (e.g. structured vs unstructured data, data size, data quality)
- Data cleaning and preprocessing techniques (e.g. missing value imputation, feature scaling, encoding categorical variables)
- Feature selection and engineering techniques (e.g. PCA, feature importance)

Setting up the Environment

- Creating AWS and Azure accounts
- Configuring the required tools and SDKs (e.g. AWS CLI, Azure CLI, Azure PowerShell)
- Understanding the infrastructure requirements for training and deploying models (e.g. EC2 instances, GPU instances, Azure ML Compute)

Model Training and Evaluation

- Choosing the right machine learning algorithm and model (e.g. regression, classification, clustering)
- Training models using AWS SageMaker and Azure Machine Learning (e.g. using built-in algorithms, custom code)
- Evaluating model performance and tuning hyperparameters (e.g. cross-validation, hyperparameter optimization)

Program Syllabus

TERM 4

Deployment AWS+Azure

Module - 3 (10 hours)

Model Deployment and Management

- Deploying trained models on AWS SageMaker and Azure Machine Learning (e.g. creating endpoints, batch inference)
- Monitoring model performance and managing versions (e.g. model drift, A/B testing)
- Integration with other services and applications (e.g. AWS Lambda, Azure Functions) techniques (e.g. PCA, feature importance)

Advanced Topics in Machine Learning on AWS and Azure

- Deep learning techniques and architectures (e.g. neural networks, convolutional neural networks, recurrent neural networks)
- Natural Language Processing (NLP) use cases (e.g. text classification, sentiment analysis, language translation)
- Understanding the costs and pricing models for machine learning on AWS and Azure (e.g. instance pricing, storage pricing, model deployment pricing)

Real-time Projects

1

Domain: BFSI

Learn and develop classification techniques for the digital transformation of banking

JPMorgan offers tax-friendly insurance choices. You can help them forecast insurance premiums. Targeted marketing using your random forest algorithm skills can help obtain better premium values.

Data Analytics, Matplotlib, Logical Regression

2

Domain: Media

Building a content recommendation model on the basis of regional viewer categorization

Netflix is a global entertainment video streaming site. They offer content in various regional languages. Build a local recommendation engine for Netflix customers residing in south Bangalore on their weekend and weekdays activities, utilizing NLP.

Data Analytics, Matplotlib, Logical Regression

3

Domain: Transportation

Reduction of waiting time via a highly precise forecasting model

Make a demand forecasting model based on specific time period rider demands. Such a model will help both riders and cab drivers to ensure the least possible waiting time. You can include measures like latitude and longitude identification.

Machine Learning, Hadoop, Time Series Analysis

4

Domain: Oil, Gas and energy

Understanding in-depth about logging while drilling (LWD) technique

Saudi Aramco company is working on the development of high-efficiency drilling models. Use the bright sides of big data analytics to identify the most cost-effective and highly productive drilling sites.

Matplotlib in Python, Big Data

Real-time Projects

5

Domain: HR

Career progression planning of employees with workforce defections & efficiency

IBM intends to boost its HR department by identifying employees' masked inconsistency. They need models to identify the graphical variations in their 14000+ employees' performances. Help them build models with your regressions and other ML abilities.

Machine Learning, Python, SQL, PySpark

6

Domain: Marketing

Descriptive study of trends and irregularities with prediction analysis for conversion

Swiggy seeks a broad marketing campaign. But they need automated keyword generation tools & proper message preparation and delivery of the same to the right audience at the right time. Help them with text analytics and NLP-based keyword research.

Exploratory Data Analysis, Big Data, NLP

7

Domain: Sales

Forecasting future sales with trends and price maximization

BMW customers can sell old vehicles, but rivals provide superior resale prices. BMW's data science-powered software will deliver the greatest market value for used vehicles based on Km travelled, daily price changes, production dates, etc. Such tasks build analytical abilities.

Scikit-learn, XG Boos, Customer Segmentation

8

Domain: Healthcare

Understanding covid-19 cases and fatality rate by time series forecasting

Samsung will launch a new healthcare app soon. The key goal of this app is an accurate human activity tracking and providing relevant health-related recommendations. Continuous analysis of a massive amount of mobile data is required for such an app.

Supervised Machine Learning, Python (Pandas Library)

Contact Us

#1090
1st floor, 18th Cross Rd, above Sangam Sweets, Sector 3, HSR Layout, Bengaluru, Karnataka 560102

Book a counselling session with expert

Book Now

Follow us on