Learnbay

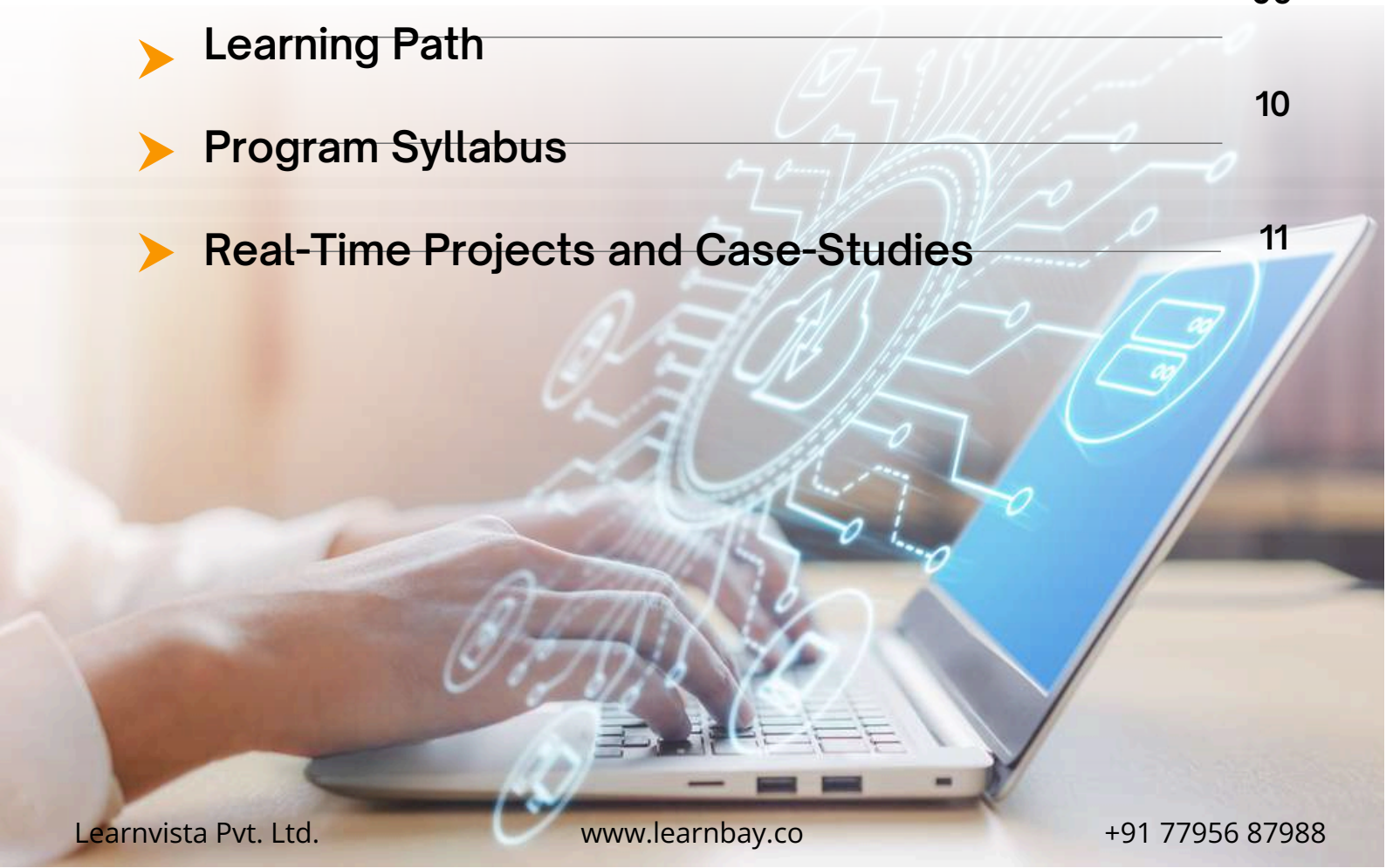
Data Science & AI Certification Program

For Managers and Leaders

- Domain Specialization
- Guaranteed Interview Calls
- Project Certification from IBM

www.learnbay.co

Context



About The Program

The Data Science & AI Program for Managers and Leaders equips professionals with essential skills in data analytics and AI, tailored for leadership roles. Focused on strategic decision-making and AI implementation, it prepares participants for advanced management positions. Additionally, our program covers **generative AI** modules, enhancing your skills in this cutting-edge field. We also incorporate ethical aspects, providing an advantage to professionals in the job market.

"We exist to provide accessible, reasonable, andindustry-relevant education that empowersIndia's workforce to grow and develop.

2cr
worth scholarships awarded

600+
professionals secured jobs after a career break

35k+
Trusted Learners

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 Details

COURSE PREREQUISITE

There are no such hard prerequisite criteria. Just the urge to learn programming and basic ideas about advanced math is enough.

PROGRAM ELIGIBILITY

- **Working professionals** having more than 1+ year(s) of experience in any domain (Technical/Non-Technical)

KEY FEATURES

- Domain Electives (BFSI, Retail, Supplychain, Marketing, Healthcare)
- Dedicated Placement Cell | 100% Guaranteed Interview calls
- Globally Recognised **Certification from IBM & Microsoft**

JOB ROLES TO TARGET

Get equipped with the industry relevant skills and aim for job roles like Data Science Manager, AI Project Manager, Business Intelligence Manager etc.

Click below

Check Eligibility

Alumni Spotlight

Shravanthi A

Data Scientist

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.

Mechanical Domain ➡ **Data Scientist @** Ford

230% Salary Hike

Preksha Mishra

Lead Data Scientist

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.

Telecom Domain ➡ **Data Scientist @** HCL

140% Salary Hike

Karan Chawala

xceedance

⬇

EXL

Data Scientist

Jaya Sinha

ADITYA BIRLA GROUP

⬇

Infosys

Senior Analyst

Shubham Dev

publicis sapient

⬇

brillio

Lead Data Analyst

Alumni Spotlight

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.

Mechanical Domain ➡ **Data Scientist @** TCS

210% Salary Hike

Saurabh Kumar
Data Scientist

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.

Mathematics Professor ➡ **Data Scientist @** Teleperformance

135% Salary Hike

Aravind

TheMathCompany ⬇ CATERPILLAR

Senior Data Scientist

Ritesh Kumar

unacademy ⬇ Capgemini

Data Scientist

Ramki

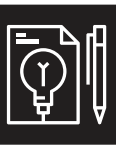
cognizant ⬇ ANBSYSTEMS energizing solutions

Data Analyst

Learnbay's ProjectLab

Choose Learnbay for your career journey because we're more than just a training provider. Our Project Innovation Lab lets you apply your skills in real-world scenarios. Get dual certifications for a competitive edge. Specialize in your desired domain. Discover how Learnbay can boost your career growth. Don't settle for less – choose Learnbay, your path to success!

1 Project Innovation Lab



Work in an industry like environment and gain practical hands-on experience of data scientist with dedicated mentors from industry.

2 Domain Electives

Enhance career prospects and excellence in your chosen field with our domain electives.

3 Dedicated Placement Cell

Experience 100% job assistance with guaranteed interview call from leading MNCs and startups globally.

4 Degree & Certification

Gain top-notch skills for a successful career through our degree and certification program

1 Project Innovation Lab

Learnbay's Project Innovation Lab replicates industry like environment for real time projects. With our **ProjectLab**, you gain real proof of hands-on experience by having your project certified by the industry.

In our ProjectLab, you work like a data scientist with dedicated project mentors from industry and get certified on capstone project.

450+
Hiring Partners

PUNE

1-1 Doubt Session

HYDERABAD

Capstone Project Certificate from IBM

35k
Trusted Learners

DELHI

BENGALURU

Project Innovation Labs Across India

2 Domain **Electives**

We offer Domain Electives for various industries including BFSI, Retail, Healthcare, Manufacturing, and Supply Chain. This ensures that you gain expertise directly relevant to your current domain.

With our program, you can make your past domain experience relevant and position yourself for a significant salary hike.

(Select any one)

BFSI

Advanced analytics for risk assessment and fraud detection in finance.

Brochure

Healthcare

Data-driven insights for improved patient care and medical research.

Brochure

Manufacturing

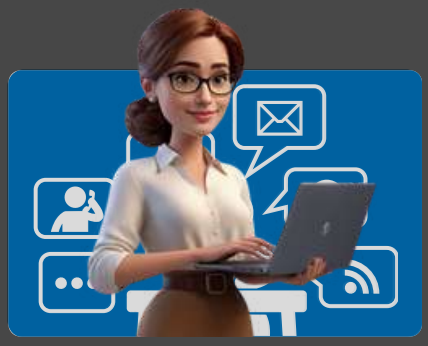
Optimising operations with data analytics and predictive maintenance solutions.

Brochure

Supplychain, E-com & Retail

Enhancing supply chain visibility and customer experience with data.

Brochure



Sales, Marketing and HR Domain

Data science can help marketers tailor their efforts to suit individual needs and preferences.

Brochure

Energy Oil and Gas

Data science can help the energy and utilities industry optimize the performance of critical assets.

Brochure

Media, Hospitality & Transportation

help businesses improve personalization, implement dynamic pricing, and perform predictive maintenance

Brochure

3 Career Service

Get 1 years of Job and Placement support

Unleash your career potential with 1 year of unlimited job access, interview support, and profile review.

1 Mock Interviews with Industry Experts

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

Resume Building Session

Craft a powerful resume showcasing your expertise in software development to stand out from the competition

4 Interview Cell

Receive 4 interview calls from a diverse pool of interested employers/recruiters.

4 Certificates

World's leading certifications

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

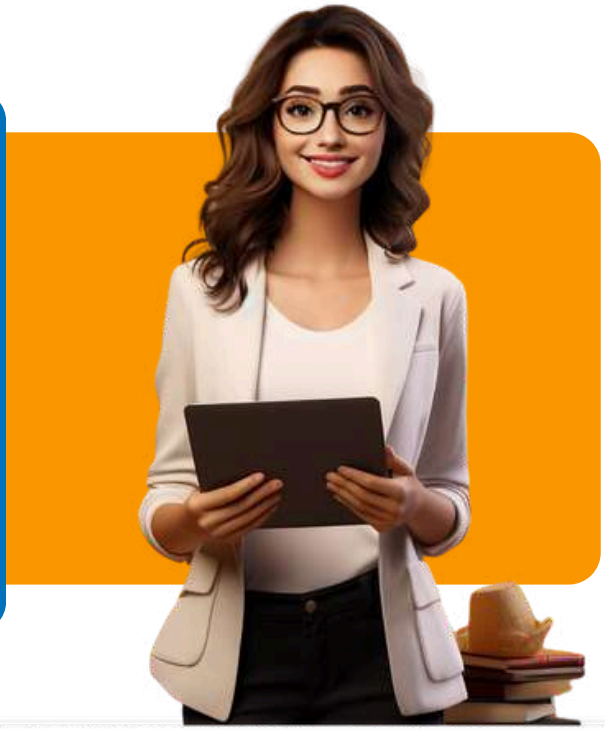
IBM Capstone Project Certificate

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

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	✓	✗
Personalised Resume Review Session	✓	✗

Program Fee & Financing



Scholarship

Scholarships are awarded based on profile review. Eligible candidates can avail upto 25% scholarship on desired courses. Click the button below to apply.

Click below

Check Scholarship Eligibility

Financing as low as

₹ 7,538/month

No Cost EMI

Razorpay

LIQUILOANS

PropelId

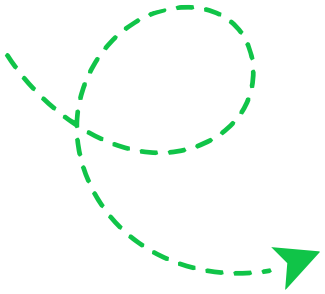
ShopSe

Program Fee

Rs. 1,15,000/-

exclusive of GST

Learning Path



06

Essentials

Project Management

05

AI Tools

Deep Learning, NLP, Deployment (AWS+GCP), CV, RL

04

Data Science Tools

SQL, MongoDB, Tableau, PowerBI, Big Data & Spark Analytics, Time Series

03

Statistics and Machine Learning

Matplotlib, Scikit-Learn, Seaborn

02

Python Programming (Basic + Advance)

Python, Anaconda, Github, Pandas

01

Cohort Orientation + Special Programming Classes

Program Syllabus

TERM 1

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. GitHub git jupyter

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

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

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

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

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

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

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

Linear Regression Model

- Comparing MAE, MSE, and RMSE. Significance of Adjusted R square.
- Overfitting and Underfitting. Bias and Variance.
- Regularization methods:
- Ridge and Lasso
- Multicollinearity, VIF. Using Python library Sklearn to create the Linear Regression Model and evaluate the model created.

Data Preprocessing

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

Regression and Classification Models

- Definition of regression, OLS Algorithm, Sum of Squares of residuals, Gradient Descent Algorithm, Cost Function
- Evaluation Metrics for Regression Model: MAE, MSE, RMSE, R Square, Adjusted R Square

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

Logistic Regression Model

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

Program Syllabus

TERM 2

Machine Learning

Module 2 (40 hours)

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

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

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

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

Program Syllabus

TERM 2

CASE STUDY

Module 2 (40 hours)

- 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

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

Power BI

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 & Spark Analytics — Module 5(14 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

Getting to know PySpark

- Pyspark Introduction
- Pyspark Environment Setup
- pySpark - Spark context
- RDD , Broadcast and Accumulator
- Sparkconf and Sparkfiles
- Spark MLlib Overview Algorithms and utilities in Spark Mlib

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

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

Project

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

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

Tools covered

TensorFlow

Program Syllabus

TERM 4

Deep Learning Using Tensorflow Module 1 (20 hours)

Introducing Recurrent Neural Networks skflow: RNNs in skflow

- Application use cases of RNN
- Manual Creation of RNN Long Short-Term Memory (LSTM) And GRU theory Restricted Boltzmann Machine(RBM)
- Autoencoders Collaborative Filtering with RBM Dimensionality Reduction with Linear Autoencoder

Understanding Of TFLearn APIs

- Getting Started With TFLearn High-
- Level API usage -Layers
- Built-in Operations
- Training and Evaluation- Customizing the Training Process
- Visualization APIs Sequential And Functional Composition Fine-tuning
- Using TensorBoard with TFLearn

Understanding Keras API for implementing Neural Networks

- Getting Started With Keras APIs Keras Model
- Sequential And Functional Model, shared layers
- Composing a Model with Keras API
- Batch Normalization
- Tensor Board With Keras
- Installing Pytorch Matrices
- Torch to NumPy Bridge
- Variables, Gradients.
- PyTorch Autograd Module
- Linear Regression With PyTorch
- Logistic Regression With Pytorch

CNN in PyTorch

- Use PyTorch to build CNN
- Build RNN with PyTorch
- LSTM in PyTorch
- LSTM from CPU to GPU in PyTorch

Real-Time Project

- SPAM Prediction using RNN
- Image Classifier using PyTorch

Tools covered

PyTorch

Keras

Program Syllabus

TERM 4

Natural Language Processing | **Module 2 (40 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

Natural Language Processing

Module 2 (40 hours)

Text Pre Processing Techniques

- Need for Pre-Processing
- Various methods to Process the Text data
- Tokenization, Challenges in Tokenization
- Stopping, Stop Word Removal

Stemming

- Stemming - Errors in Stemming
- Types of Stemming Algorithms - Table
- Lookup Approach
- N-Gram Stemmers

CASE STUDY

- **Sentiment analysis for Twitter, web articles**
- **Movie Review Prediction**
- **Summarization of Restaurant Reviews**
- **Topic Modelling & Dirchlett Distributions**
- **Introduction to Topic Modelling**
- **Latent Dirchlett Allocation**
- **Advanced Text Analytics & NLP**
- **Introduction to Natural Language Toolkit**
- **POS Tagging**
- **NER (Named Entity recognition)**

Program Syllabus

TERM 4

Computer Vision | **Module 3(24 hours)**

Computer Vision overview

- Historical Perspective
- Introduction to the four R's of Computer Vision
- OpenCV Installation
- Python API Drawing shapes
- Image Processing
- Image Rotation and Thresholding

Image Processing

- Histogram equalization
- Thresholding and Convolution
- Sharpening and edge detection
- Morphological transformations
- Image pyramid
- Contour properties
- Circle detection
- Line detection
- Watershed segmentation

Project

- AI-Based Live Face Identification System for Crowd
- Single Shot MultiBox Detector,
- Object Localization
- Find an object in an image

Image Filtering

- Gaussian Blur
- Median Blur Feature Detection - Canny Edge Detector
- Use of Neural Network in CV
- Multi-Layer Perceptron

Image Classification and segmentation

- Data-Driven approach
- K-nearest Neighbor
- Linear Classification
- Contours and segmentation

Project

- The Problem of Scale and Shape
- Haarcascade - face and eye detection

Real time use cases:

- Single Shot MultiBox Detector
- Object Localization
- How would you find an object in an image
- The Problem of Scale and Shape SSD in Tensorflow
- Haar cascade - face and eye detection

Program Syllabus

TERM 4

Reinforcement Learning | **Module 4(14 hours)**

What is Reinforcement Learning - Basics

- Setting up Environment & Installing OpenAI Gym.
- OpenAI Gym Basics.
- Terminology & Environment.
- Dynamic Programming - Prediction, Control, and Value Approximation.
- Building Blocks of Reinforcement Learning
- OpenAI Gym Tutorial Random Search
- Markov Decision Processes
- Monte Carlo Methods

Projects and Case Studies

- Solving Taxi Environment
- Solving Frozen Lake Environment
- Reward Discounting

Approximation Methods for Reinforcement Learning

- RBF Networks with CartPole
- TD Lambda and Policy Gradient Algorithms.
- Temporal difference learning. N-Step Methods
- TD lambda, Policy Gradient Methods Policy Gradient in TensorFlow for CartPole. Mountain Car Continuous using Tensorflow

Important:

- Deep Q-Learning Techniques
- Deep Q-Learning in Tensorflow for CartPole

Program Syllabus

TERM 4

Model Training & Deployment using (AWS GCP) **Module 5 (10 hours)**

AWS (Amazon Web Services)

- Deployment Strategies
- Automations
- Monitoring and Logging
- Communication and Collaboration

GCP (Google Cloud Platform)

- GCP Development Tools - Cloud SDK, Repositories, Plugins
- Deployment Manager and Cloud Endpoint

Deploying Machine Learning Model

- Deploying Models, Understanding training graphs and serving graphs
- Check and adjust model size
- Build an optimal prediction graph
- Creating input function
- Creating a model version
- Getting Online Prediction

Introduction to AWS and GCP Cloud ML Engine

- CloudML Engine & AWS in Machine Learning WorkFlow
- Components of AWS & Cloud ML Engine
- GCP and AWS Console.
- gcloud command-line tool and Rest API

Training Machine Learning Model

- Developing a trained model application
- Running and monitoring a machine learning model
- Using hyperparameter tuning
- Using GPUs for training models in the cloud

Tools covered

aws

Google Cloud

Program Syllabus

TERM 5

Project Management Module | **Essentials (70 hours)**

Jira process

- Agile Delivery and Scrum DevOps
- Project Management
- Release Management Process
- Service Now
- Meetings/Emailing
- Communication with various workstreams
- Change Management
- Resource Management
- Stakeholder Management
- Risk analysis to improve outcomes
- **Risk Management**
- RAID log
- Realistic time estimates
- Project Charter
- Co-create a project task outline and schedule
- Status Tracking
- Project Management
- Agile Project Management
- **Project Management Cycle**

Agile

- Agile Delivery and Scrum
- Agile & Scrum in a nutshell
- Lifecycle of a Scrum-based project
- Scrum Roles
- Scrum: Sprint Lifecycle (Ceremonies)
- Scrum Artefacts
- Business Requirements

PM approaches for Technical Projects

- Managing Technical solutions
- PM Tips
- Plan on a Page (POAP)
- Roadmap Project Planning
- Detailed Project plan
- Project Management Tool (MPP)
- Case studies
- Certification guidance

JIRA Process

- What & Why Jira
- Delivery Process enabling
- Getting access & requesting a new projects on Jira
- Adding team members to your Jira Project
- Navigating Jira
- Jira Issue Types
- Jira Training assets
- Jira Reports

Program Syllabus

BONUS

AI Generative Tools and Future Trends

Emerging Trends in AI and Generative Modeling

- Exploring other AI generative tools beyond ChatGPT and DALL·E
- Overview of Midjourney
- Discussion on future trends and advancements in AI generative tools
- Open-ended project and/or presentation on a selected topic, incorporating learned concepts

Natural Language Processing and ChatGPT

- Introduction to natural language processing techniques
- Understanding ChatGPT and its architecture
  Hands-on exercises using ChatGPT for text generation
- Fine-tuning ChatGPT for specific applications

Midjourney Bing ChatGPT DALL·E Bard

DALL·E: Image Generation with AI

Exploring image generation using DALL·E

- Hands-on exercises for creating unique images with DALL·E
- Ethical considerations and limitations of AI-generated images

Graph Neural Networks (GNN) for Data Analysis

- Overview of Graph Neural Networks (GNN) and their applications
- Hands-on exercises using GNN for tasks such as node classification and link prediction
- Case studies on real-world applications of GNN in data science

Program Syllabus

TEBROMN 1U

Python Bootcamp for AI

- Python Essentials: Syntax, Data Types, and
- Variables Flow Control: Conditionals and Loops

Build Your Interview Assistant

- Project Overview: Interview Automation Bot
- Components & Architecture
- Natural Language Models (LLMs): Introduction and Uses

Large Language Models (LLM

- Historical Overview of NLP: From Rule-Based Systems to Machine Learning.
- Evolution of Neural Network Architectures in NLP.

Visual AI for eCommerce

- Introduction: Digital Transformation for Offline Businesses
- Multimodal Models: DALL-E and Beyond
- Style & Photography Principles for Visual AI

Intelligent News Aggregator

- Project Outline: Personalized News Recommendation
- GPT-3 & Copilot for Code Automation

Program Syllabus

BONUS

Customer Support Bot - HelpMate Pro

- Project Introduction and Components
- Embeddings vs Fine-Tuning: When and How
- Semantic Search in Customer Service

Knowledge Discovery Bot

- Project Overview and Architecture
- LangChain Tools and Concepts

Azure OpenAI Integration

- OpenAI on Azure: Services and Scalability
- Revisiting HelpMate Pro: Scaling Strategy

The Future & Ethics of Generative AI

- Responsible AI: Bias and Fairness
- Future Trends: Multimodal Models and RLHF
- Closing Remarks
- Assessment: MCQ

Capstone Project (3 Weeks)

- Building an Integrated Prompt Engineering Solution
- Project Submission and Peer Review

The **IBM** exam will be conducted for all the modules after completion of the course

Real-time Projects

13 hours

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.

python R

17 hours

Descriptive study of trends and irregularities with prediction analysis for conversion.

Swiggy seeks a broad marketing campaign. They also require proper message preparation and delivery of the same to the right audience at the right time. You can help them with text analytics and NLP-based keyword research.

sas MATLAB

21 hours

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.

tableau python

15 hours

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.

pandas TensorFlow 2.0

Real-time Projects

13 hours

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.

python Apache Spark

17 hours

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.

TensorFlow 2.0 NumPy

21 hours

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.

NumPy pandas

15 hours

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.

MATLAB python

Real-time Projects

13 hours

Churn forecasting for the telecom industry using R programming with ML

The goal of this project is to design a precise customer churn prediction model. Based on the same, Jio can identify the exact reason for customer dissatisfaction and work accordingly.

python MATLAB

17 hours

Recommendation system with customer lifetime value analysis (CLV)

Amazon wants to find the most successful electronics. Live consumer reviews are needed. Using data visualisation, help regenerate consumer insights from ongoing and current reviews.

pandas NumPy

21 hours

Condition-based preventative maintenance and fault prediction in depth

This project helped BOSCH to predict their internal failures by production line dataset analysis. But still, they are struggling to predict automated faults in their assembly stage. Help them by building more advanced predictive models for assembly stage monitoring.

TensorFlow 2.0 pandas

15 hours

Automated inventory monitoring for supportable supply chain management

An automated inventory management system will keep track of stock levels and upcoming orders. In addition, you can contribute to DataCo's intelligent supply chain software generation project by using ML algorithms and R programming skills.

Power BI python

Contact Us

Click here to whatsapp

or call us at
+91 77956 87988

www.learnbay.co