

Data Science and GenAI

Master Certification Program

*Get Certified on
Real Projects from AI Startups*

In collaboration with



*Curriculum inclusive of
GenAI and Agentic AI*



Table Of Content

1	How to make an Impact
2	Program Summary
3	About Course
4	Who is this program for?
5	Why choose Learnbay?
6	Others Vs Learnbay
7	Alumni Spotlight
8	Certification
9	Fee & Batch Details
10	Program Curriculum

How to make an
impact with



*This program is designed for **working professionals** who want to master full-stack **Data Science and Generative AI** in depth to become an **AI engineer in top product-based MNCs**.*



Project Experience Certificate

Earn project experience certificate directly from **GenAI startups**



Program Certification

Earn a completion certificate from IBM & Microsoft



Dedicated Career Assistance

Personalized 1:1 support to accelerate your career growth.



GenAI and Agentic AI

Learn LangChain, AutoGPT, BabyAGI, CrewAI with hands-on projects.

Program Summary

1

Training Mode

100% Live Online & Hybrid
(Online + Classroom)

2

Program Faculty

Renowned industry experts
from top global MNCs

3

Program Duration

13 Months Program Weekday
and Weekend Batch

4

Certification

IBM &  **Microsoft**



Project Based Learning

Exclusive projects with real AI startups through the AI Co-Lab



What is AI Co-Lab Experience?

AI Colab is a unique platform where learners work directly on real projects from AI startups, solving actual business challenges in GenAI & Agentic AI. You don't just learn—you co-create with startups.



Who will guide me during these projects?

Each project is supported by startup founders, industry experts, and Learnbay mentors with 1:1 guidance, feedback



Why AI Co-Lab Startup Certification matters?

A project certificate is more valuable than a course completion certificate as it showcases your expertise and practical abilities.

About Course

This program is designed for working professionals who want to gain deep knowledge of full-stack Data Science and AI, including advanced concepts. It uses a **project-based learning approach**, where you'll collaborate with industry experts and have support from a dedicated project mentor.

The curriculum covers **Generative AI (GenAI)** and offers domain specialization, allowing you to work on capstone projects in areas like **sales, retail, supply chain, finance, and HR**.

Our Commitment

"We are dedicated to delivering accessible and industry-relevant education that empowers India's workforce to grow and succeed."



Beyond classroom projects, you'll co-create solutions with AI startups through our exclusive AI Co-Lab initiative.

Furthermore, our comprehensive career support services include **interview preparation, resume building, and job placement assistance**, all designed to help you smoothly transition into leadership roles in data science and AI.



100%
Assured Interview



Learners working at
350+ companies



7+
Centers Across India



40k+
Professionals Upskilled



70%

of companies are likely to adopt GenAI in next 5 years, to

- enhance efficiency,
- automating tasks,
- improving decision-making for better project outcomes.

***By integrating GenAI into our program**, we ensure that our learners are well-prepared to lead and innovate in their respective fields.

Program **Eligibility**

**Working
professional
with min
1+ Year of
Experience**

Professionals with a minimum of **1 year** of experience in any domain. Prior knowledge of Programming is **not mandatory** as we provide **special bootcamp for non programmers** with this course.



Executives (IT/Non-IT)

Important Note: This program is **not for freshers and students**. If you are a fresher, please refer our Data Science and AI program for freshers.

Program Outcome: What's in it for you?



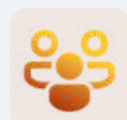
Hands-On Experience With AI Co-lab

Work directly on real startup challenges & get certified.



Domain-Specific Mastery

With a domain-specialized data science course, your previous work experience becomes relevant, making it easier to switch careers. You can work on projects from various fields like **BFSI, Retail, Healthcare, HR, Marketing**.




Learn Full Stack Data Science & Gen AI in Depth


Learn Data science and AI in depth with specialization in Advance Deep Learning, NLP, Computer Vision and GenAI. This program will help you to become **AI engineer in top product based MNCs**.



Why choose Learnbay?

A unique program for
Working Professionals!

 www.learnbay.co

 77956 87988

AI Co-Lab Experience



A unique feature where learners work directly on real projects from AI startups, solving actual business challenges in GenAI & Agentic AI. You don't just learn—you co-create with startups.

Gain hands-on experience with GenAI & Agentic AI



Hugging Face



End to End Support

Each project is supported by **startup founders, industry experts, and Learnbay mentors.**

You get **1:1 guidance, feedback, and code reviews**—like working in a real company environment.



Earn **industry-recognized startup certifications**. You can highlight your work on real AI applications

You'll collaborate in small teams, just like in startups. The process includes:

- **Startup briefing**
- **Problem scoping**
- **Coding & implementation**
- **Mentor feedback**
- **Final presentation & certification**

How is this different from regular course projects?

Unlike typical academic case studies, these **projects are live startup problems**. You'll see your work impact real users, products, and decisions—something no classroom project can match.

How will this help me as a working professional?

You'll gain **hands-on experience with GenAI & Agentic AI**, add real startup certifications to your resume, and be able to talk about actual business outcomes in interviews, giving you a huge career edge.

What kind of companies provide these projects?

Startups and growing enterprises that are building AI-first products across industries like **FinTech, EdTech, Healthcare, Retail**, and more.

Why is a project certificate from an AI startup better than a course completion certificate?

Course certificates show you've learned. **Startup project certificates** prove you've delivered. Recruiters value impact-driven proof more than just attendance.

How will this accelerate my career growth?

By completing AI Colab projects, you'll:

- Gain practical AI experience
- Stand out in interviews
- Get startup-certified proof of work
- Build a portfolio recruiters love

Do I need to be an advanced programmer to join?

No. Projects are designed for working professionals from **tech and non-tech backgrounds**. You'll be trained on tools and supported by mentors throughout the journey.

1. Project Based Learning

Project based learning gives you **hands-on experience** to apply concepts to real-world problems, sharpening your skills for practical use. Unlike traditional methods, our Project based learning approach prepares you for practical, real-life challenges.



**Live Projects from
industry**



**Dedicated Project
Mentor**

Important Note: Our program includes capstone, mini, guided, and domain projects. You'll have a dedicated mentor to guide and evaluate your project work. You can book **unlimited 1-on-1 sessions** with your project mentor for personalized support. Please refer syllabus section for more details.

2. Domain Specialization

Make your previous experience count

Select your domain electives and engage in live projects across various industries to gain hands-on experience. This practical approach will deepen your understanding of real-world challenges and enhance your expertise.



Retail



Technology



Consulting



Manufacturing



BFSI



Healthcare

3. AI Startup Project Certification

Earn certificates directly from companies you work on across different domains. These project certificates are **more valuable** than course completion certificates because they show your real skills and practical experience.

Get **Project Certification**
from AI/GenAI Startups.

ElevonAi

 **Outspokn**



Fermion

CodeMYX & more

You will work on projects like:

AI-powered content creation (text, images, video, audio), customer support through **chatbots and virtual assistants**, software development tools to **assist coding, personal productivity agents** for task automation, and data analysis for insights in fields such as finance, healthcare, and marketing

4. 1:1 Doubt & Mentorship

You can book **personalized doubt classes** for modules or projects with experts.

These tailored sessions will help you understand concepts better and **address any questions or challenges** you might have.

Important Note: Schedule your sessions from **10 AM - 6 PM (Mon-Sat)** to secure your preferred time slot.

5. Training Mode

You can choose from two flexible training modes according to suit your needs and preferences:

New



**100% Live
online classes**



**Hybrid
classes**

Live Online Mode: In this mode, you attend all sessions and work on projects live online. The classes are instructor-led, allowing you to ask questions in real-time. You'll also have access to online project mentors and career mentors for 1-1 sessions.

Hybrid Mode: You can attend sessions/classes online, but for live project work, you have the option to participate in [physical classrooms at our Project Innovation Labs](#) in cities like *Bangalore, Pune, Delhi, Hyderabad, Chennai, and Kolkata*.

6. Dedicated Placement assistance



Resume Optimization

Expert assistance to enhance your professional resume



Interview Preparation

Mock interviews to improve your performance



Interview Opportunities

Scheduled interviews with potential employers



Career Counseling

Professional guidance for your career advancement

7. Inclusive of GenAI and Agentic AI

40% job growth in GenAI & Agentic AI roles by 2027

Secure future-ready career opportunities.

\$1.3 trillion GenAI market by 2030

Tap into high-paying roles in AI-powered domains.

90% of businesses to adopt GenAI by 2026

Gain a competitive edge in AI-driven industries.

Agentic AI improves decision accuracy by 85%

Enhance project execution and business strategies.


Automate 60% of repetitive tasks

Use GenAI for emails, reports, and content, and Agentic AI for workflow automation.

80% of professionals use GenAI daily

Leverage AI tools to boost efficiency and productivity at work.

Others Vs Learnbay

Features	Learnbay's MasterTrack Certification	Other Certification
Access to Live Classes	✓ 3-Year Flexi Pass for Live Interactive Classes	✗ Usually Limited (6-12 months)
Class Type	✓ Live, Mentor-Led, Interactive	✗ Mostly Recorded Sessions
Domain Specializations	✓ BFSI, Healthcare, Supply Chain, Manufacturing, HR, Project Management	✗ Limited or Generalized
 AI Startup Project Certification	✓ Issued by Industry Partners	✗ No project certificate
Learning Mode	✓ Hybrid (Online + Classroom Project Sessions)	✗ Primarily Online
Project Mentorship	✓ Dedicated Project Mentor	✗ General Doubt Sessions or Not at all

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 @



 **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**



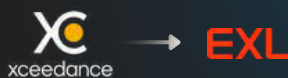
Data Scientist @



 **140%**
Salary Hike



Karan Chawala



Data Scientist



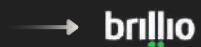
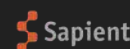
Jaya Sinha



Senior Analyst



Shubham Dev



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 @



 **230%**
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**

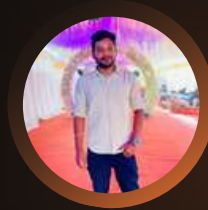
Data Scientist



Ritesh Kumar

 **unacademy** → **Capgemini**

Senior Analyst

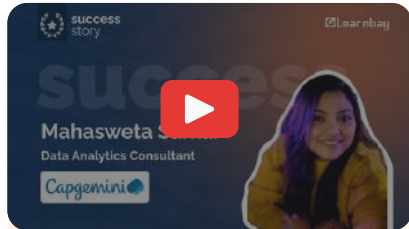


Ramki

 **cognizant** →  **ANBSYSTEMS**
#ONYX company

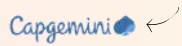
Data Analyst

Success Story



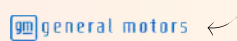
Mahasweta Sarkar

Data Analytics Consultant



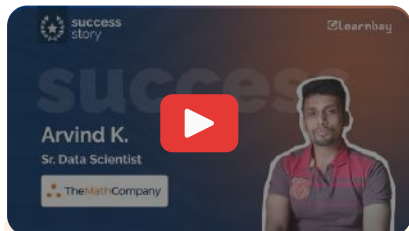
Manoj Kuna

Data Analysis Engineer



Bhavin Shah

Data Analyst



Arvind K.

Sr. Data Scientist



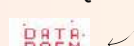
Nandini Devi Muthu

Data Analyst



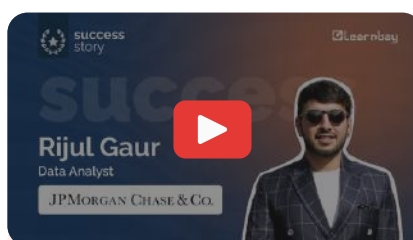
Rajashree B

Data Quality Analyst



Pinky Bhansali

Process Lead



Rahul Gaur

Data Analyst



Jaya Sinha

Data Scientist



Get certified

and accelerate your career growth

IBM Course Completion Certificate



Achieve an **industry-recognized IBM certification**, showcasing your proficiency in **data science and AI**.

Microsoft Certificate in Azure AI Fundamentals



Gain an **internationally recognized Microsoft certification**, proving your expertise in **cloud technologies**.

Add On

AI Co-Lab Project Certificate



This certification boosts your credibility in the **IT sector** and enhances your **career prospects**.

Certification from IBM in Generative AI



Hands-on exposure to **industry-relevant projects** using IBM Watsonx and leading **Generative AI tools**.

Program Fee

Live online classes

- ✓ Live online interactive sessions
- ✓ 1:1 online Doubt Session with experts
- ✓ Virtual Mock interviews
- ✓ Online Capstone projects

Program Fee

₹ 1,59,000/- +18% GST

EMI starts at

₹ 13,250/-

Payment Plan:

No Cost EMI Available
For 6, 9 and 12 Months

Popular Tools Covered:

GenAI:



ChatGPT



DALL·E

Gemini



Midjourney



Claude

Agentic AI:



LangChain

crewai


AUTOGPT



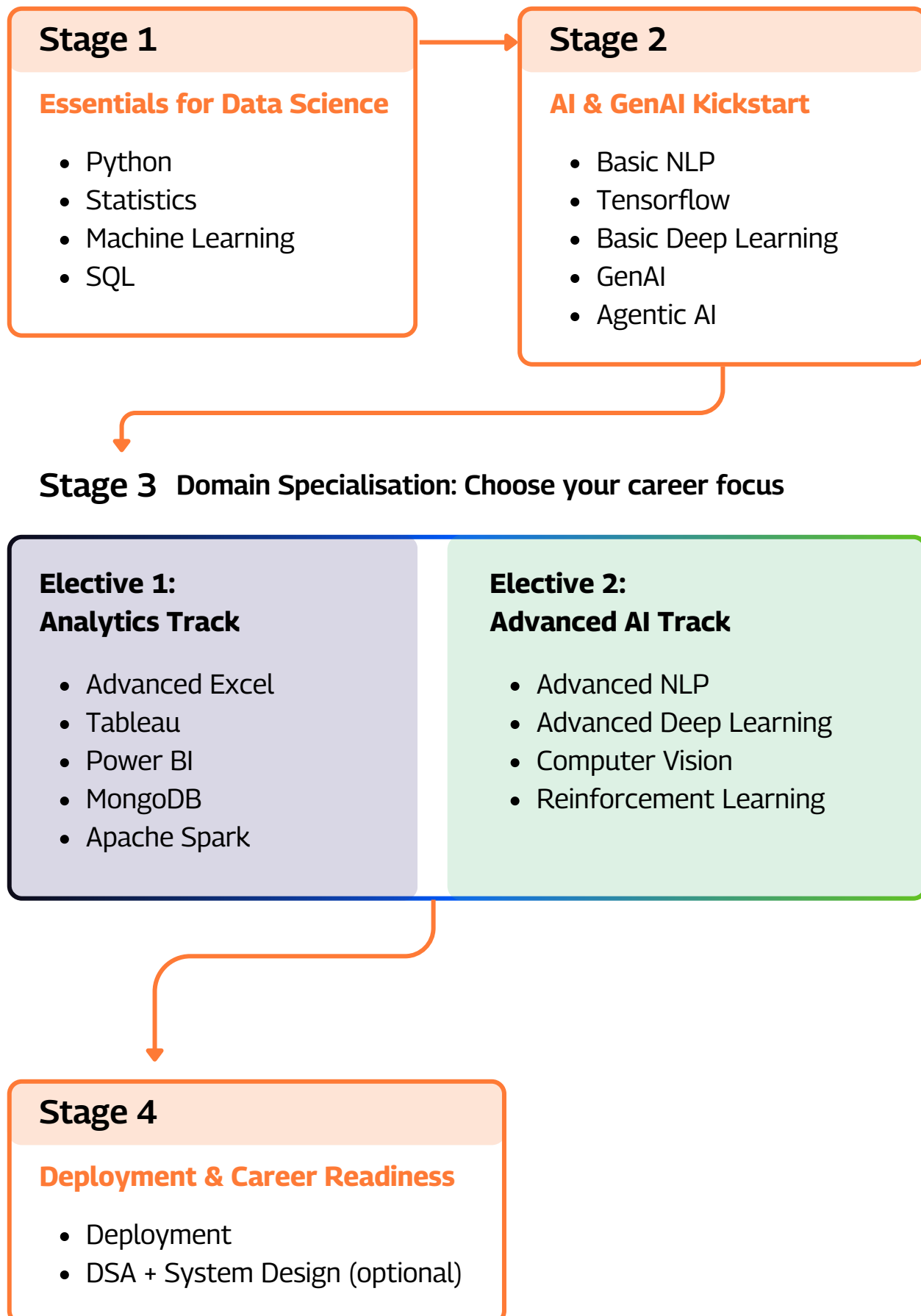
Program Curriculum

A unique program for
Working Professionals!

 www.learnbay.co

 77956 87988

Journey to **Upskilling**

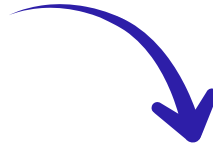


How does **AI Co-Lab** Works?

01

Team Project Assignment

Get assigned to real startup projects in small learner teams, just like industry squads.



02

Work on Real Problem Statements

Solve actual business challenges provided by AI-first startups, not dummy datasets.



03

Guided by Startup Founders & Industry Experts

Receive mentorship and insights directly from AI startup founders and domain leaders.



04

Build with GenAI & Agentic AI + Earn Certification

Use cutting-edge GenAI & Agentic AI tools to deliver solutions. Get certified by the startup itself for your work.

Journey to Upskilling

Module 0 8 hrs

Python Bootcamp for Non-Programmers

These sessions are recommended for non programmers to learn basics of programming.



PYTHON

Term 1 40 hrs

Python for Data Science

We will cover Python basics and data analytics using popular libraries like NumPy, Pandas, Matplotlib, and Seaborn.



PLOTLY



SCIKIT LEARN



NUMPY



PANDAS



MATPLOTLIB



SEABORN

Term 2 70 hrs

Statistics and Machine Learning + GenAI

We will cover statistical methods and explore advanced topics like regression analysis, hypothesis testing, EDA and machine learning algorithms.



STATISTICS



MACHINE LEARNING



GENAI

Term 3 50 hrs

Generative AI and Agentic AI

Master GenAI for content generation, data analysis, and NLP, and Agentic AI for automation, decision-making, and workflow optimization.



ChatGPT



DALL-E



AUTOgpt



LangChain

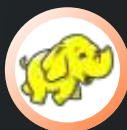
Term 4 60 hrs

Data Visualization & Data Analysis

We will cover data visualization and data analysis including interactive dashboards, statistical analysis, data mining, and predictive modeling.



POWERBI



HADOOP



APACHE SPARK



TABLEAU



SQL

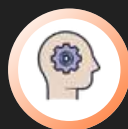
Term 5 70 hrs

AI Tools + Deployment

We will cover AI tools fundamentals and advanced features including machine learning frameworks, natural language processing, and computer vision applications.



TENSORFLOW



DEEP
LEARNING



NLP

Term 6 30 hrs

Excel

We will cover Excel fundamentals and advanced features: formulas, functions, pivot tables, and data visualization.



EXCEL

Term 7 30 hrs

DSA + System Design

We will cover DSA using popular languages like Python and Java. You'll understand concepts such as arrays, linked lists, trees, sorting, and searching algorithms.



JAVA



PYTHON

Term 8 30 hrs

Project Management

We will cover project management fundamentals and methodologies like Agile, Scrum, and Waterfall. You'll work on several guided projects to apply these skills in real-world scenarios.



EXCEL

370+ hours
of Learning

34+
Practical Projects

Triple
Certification

Module 0 Python Bootcamp for Non-Programmers

The Python Bootcamp is designed for **professionals who are new to programming**. It covers basics at a slow pace and you'll get **1-1 doubt sessions and special classes**. These sessions are recommended for non programmers. Programmers may skip this module

Topics Covered

Installation & Setup

- Installing Python and setting up IDEs (Jupyter, VSCode).
- Overview of the Python environment and running basic scripts.
- Introduction to Jupyter Notebook for interactive coding.

Basic Syntax & I/O

- Understanding variables, data types (integers, floats, strings, booleans).
- Input and output operations using the `input()` and `print()` functions.
- Performing basic arithmetic operations.

Control Structures

- Conditional logic using `if`, `else`, and `elif` statements.
- Looping with `for` and `while` loops to iterate over data.

Basic Data Structures

- Introduction to lists, tuples, and dictionaries.
- Accessing, modifying, and `iterating` through collections.
- Understanding how to use indexing and slicing with lists and tuples

MINI PROJECT

DURATION 1.5 HRS

FOR BEGINNERS

Simple Calculator Application

Description: Create a console-based calculator that takes user input for two numbers and performs mathematical operations (add, subtract, multiply, divide)

Tools/Functions covered: Python basics (Control structures, basic I/O)

TERM 1 Python for Data Science

We will cover Python basics and data analytics using popular libraries like **NumPy, Pandas, Matplotlib, and Seaborn**. You'll work on several **guided projects** in python and gain hands-on experience throughout this module.

Module Duration:

40-50 hrs of live Instructor-led classes.

5 hrs of **on demand 1-1 session** for topics where you need help.

Number of Projects with dedicated mentor

4 practice projects

2 python projects on real data

Section 1 - Core Python Basics

• Python Environment Setup

- Setting up Python environments using Anaconda and **Jupyter**.
- Installing and managing virtual environments using conda/pip.

• Python Syntax and Code Structure

- Writing and running Python scripts.
- Best practices for writing clean and readable code (PEP 8).
- Understanding Python functions and scopes.

• Data Types and Control Flow

- Primitive data types: integers, floats, strings, and booleans.
- Control flow statements (if, else, elif) for decision making.
- Loops: **for, while, do while, break, continue, pass**

• Data Structures

- Introduction to basic data structures (lists, tuples, dictionaries).
- Manipulating collections: Adding, removing, and modifying elements.
- List comprehension and dictionary comprehension for concise code.

• Functions And Modules

- Introduction To Functions Defining & Calling Functions
- **Functions With Multiple Arguments**
- Anonymous Functions - **Lambda** Using Built-In Modules,
- User-Defined Modules, Module Namespaces, Iterators And Generators

GROUP PROJECT

DURATION 2 HRS

BFSI DOMAIN

Banking System Simulation

Description: Simulate a simple banking system where users can create accounts, deposit, withdraw, and check balances.

Tools/Functions covered: Python (Data structures, functions)

MINI PROJECT

DURATION 1.5 HRS

FINANCE DOMAIN

Expense Tracker

Description: Build a simple expense tracker where users can record and manage their daily expenses.

Functions used: Python data structures, Functions.

Section 2 - Data analytics Using Python

- **File I/O And Exceptional Handling and Regular Expression**

- open Function, file Object Attributes
- `close()` Method, `Read`, `write`, `seek`.
- Exception Handling, try-finally Clause
- **Regular Expression**- Search and Replace
- Regular Expression Modifiers
- Regular Expression Pattern

- **Basic Web Scraping for Data Analytics**

- Definition and use cases in data analytics
- Tools for scraping: requests and `BeautifulSoup`
- Parsing HTML content using BeautifulSoup
- Finding elements by `tag`, `class`, or `attributes`
- Extracting text, links, and attributes from elements

- **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
- **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 Visualization using Matplotlib**
 - Introduction, **plot()**, Controlling Line Properties, Subplot with Functional Method, Multiple Plot, Working with Multiple Figures, Histograms
- **Data Visualization using Seaborn**
 - Introduction to **Seaborn** and Visualizing statistical relationships , Import and Prepare data. Plotting with categorical data and Visualizing linear relationships.

PORTFOLIO PROJECT

DURATION 2 HRS

HR DOMAIN

Employee Attrition Analysis

Description: Analyze employee attrition (churn) to identify patterns related to factors like salary, job satisfaction, department, and years of service.

Tools/Functions Covered: Pandas, Numpy and Matplotlib

DOMAIN PROJECT

DURATION 2 HRS

SALES DOMAIN

Sales Performance Dashboard

Description: Create a sales performance dashboard to track sales revenue, number of units sold, and performance by region, product, or salesperson.

Tools/Functions Covered: Pandas, Numpy and Matplotlib

DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

FINANCE DOMAIN

Stock Price Analysis with Web Scraping

Description: Use web scraping to collect real-time stock price data from a financial website. Analyze stock prices over time to identify trends, calculate returns, and visualize stock volatility.



Seaborn



pandas



Web Scraper

DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

BFSI DOMAIN

Customer Transaction Analysis

Description: In this project, you'll analyze customer transaction data from a bank to gain insights into customer behavior. You will use only Pandas for data manipulation and Seaborn for data visualization.



Seaborn



pandas

TERM 2

Statistics and Machine Learning + GenAI

We will cover statistical methods and explore advanced topics like **regression analysis, hypothesis testing, EDA** and **machine learning algorithms**. Practice all the topics with real time projects and case studies.

Module Duration:

70-75 hrs of live Instructor-led classes.

5 hrs of **on demand** 1-1 session for topics where you need help.

Number of Projects with dedicated mentor

6 practice projects

2 python projects on real data

Section 1 - Statistics

• 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

• 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

• 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

• 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

GROUP PROJECT

DURATION 1.5 HRS

Exam Score Statistical Analysis

Description: Perform descriptive and inferential statistics on a sample of student exam scores to understand performance trends.

Tools covered: Sampling methods, variable types, measures of central tendency and dispersion, five-point summary, box plots, outlier detection using IQR and Z-score methods, confidence intervals, hypothesis testing errors.

MINI PROJECT

DURATION 2 HRS

MANUFACTURING DOMAIN

Probability Analysis in Quality Control

Description: Analyze the probability of defects in a manufacturing process using probability theories and Bayes' Theorem.

Tools covered: Probability distribution and cumulative distribution functions, conditional probability, Bayes' Theorem, mutually exclusive and joint events, dependent and independent events.

Section 2 - Advance Statistics

• 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

- **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

- **EDA**

- Understand EDA's role in data science workflow.
- Load and inspect data using Pandas.
- Using Z-scores to Find Outliers.
- Clean data by handling missing values and outliers.
- Perform feature engineering for better modeling
- Finding and Dealing with Missing Values.
- Bivariate Analysis, Scatter Plots and Heatmaps.
- Summarize data with descriptive statistics.
- Visualize patterns using Matplotlib and Seaborn.
- Introduction to Multivariate Analysis

MINI PROJECT

DURATION 2 HRS

ECOMMERCE DOMAIN

Hypothesis Testing on Website Conversion Rates

Description: Test if a new website design leads to a significantly higher conversion rate compared to the old design.

Tools covered: Binomial distribution assumptions, Z-test, P-value method, Type I and II errors, one-tailed test using SciPy and Pandas.

RESEARCH PROJECT

DURATION 2 HRS

Linear Algebra in Image Transformation

Description: Apply matrix operations to perform transformations on images, such as rotations and scaling.

Tools/Libraries: NumPy matrices, matrix multiplication, matrix transpose, identity matrix, determinant, dot product.

MINI PROJECT

DURATION 2 HRS

EDA and Data Cleaning on Titanic Dataset

Description: Perform exploratory data analysis and data cleaning on the Titanic dataset to prepare it for modeling.

Tools covered: Data wrangling with Pandas, handling missing values, outlier detection using Z-scores, feature engineering, bivariate and multivariate analysis, data visualization using Matplotlib and Seaborn.

Section 3 - Machine Learning

- **Machine Learning Introduction**
 - 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**
 - Encoding the data: Definition, Methods: OneHot Encoding, Mean Encoding, Label Encoding, Target Guided Ordinal Encoding
 - **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
- **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

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Customer Churn Prediction with Logistic Regression

Description: Predict whether customers will leave a service using logistic regression.

Tools covered: Data encoding (OneHot, Label Encoding), handling missing values (MCAR, MAR, MNAR), outlier treatment (IQR, Z-score methods), feature scaling, logistic regression with scikit-learn, evaluation metrics

RESEARCH PROJECT

DURATION 2 HRS

Handwritten Digit Classification using K-Nearest Neighbors

Description: Classify handwritten digits into categories using the KNN algorithm.

Tools covered: Data preprocessing, KNN implementation with scikit-learn, distance metrics (Euclidean, Manhattan), feature scaling, evaluation with multi-class confusion matrix.

Section 4 - Advance Machine Learning Concepts

- **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

- **Random Forest Model**

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

- **Hyperparameter Tuning**

- GridSearchCV, Variable Importance.
- Using Python library Sklearn to create the Random Forest Model and

evaluate the model created.

- **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

- **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

- **XGBoost Model**

- Definition and advantages; enhances gradient boosting with regularization, handling missing values, and parallel processing.
- Implementing XGBoost in Python for model creation and evaluation; hyperparameter tuning using GridSearchCV.

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Credit Risk Prediction with Decision Trees and Random Forests

Description: Predict credit default risk using Decision Tree and Random Forest models, enhancing performance through hyperparameter tuning.

Tools covered: Scikit-learn Decision Trees, Random Forests, GridSearchCV for hyperparameter tuning, K-Fold Cross-Validation.

RESEARCH PROJECT

DURATION 2 HRS

Customer Segmentation using K-Means Clustering and PCA

Description: Cluster customers based on purchasing behavior using K-Means and Hierarchical Clustering, incorporating dimensionality reduction with PCA.

Tools covered: Scikit-learn K-Means Clustering, model evaluation.

GROUP PROJECT

DURATION 3 HRS

Predictive Maintenance for Industrial Machinery

Description: Develop a predictive maintenance model using machine learning and statistical analysis to predict equipment failures. Combine EDA, statistical methods, and supervised learning models for failure prediction.

XGBoost



Matplotlib



Numpy



pandas

Includes AI Co-Lab real startup challenges. 

DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

Customer Segmentation and Churn Prediction

Description: Perform customer segmentation using clustering techniques and predict customer churn with logistic regression. Use statistics for EDA and apply machine learning for churn classification.



Matplotlib



Seaborn



pandas



scikit-learn



statsmodels

Master advance GenAI tools like **TensorFlow, PyTorch, OpenAI APIs, and LangChain**, frameworks, and **real-world applications** to lead AI-driven projects confidently.

**50 Hours instructor-led
6-10 real time GenAI &
Agentic AI Projects**

Section 1: Generative AI

Fundamentals of Generative AI

- Overview of **generative models** and their real-world applications, highlighting differences from traditional AI.

Use Cases of GenAI in Various Industries and Domains

- **Finance:** Fraud detection, automated report generation, and investment insights.
- **Healthcare:** Drug discovery, medical image analysis, and personalized treatment,
- **Retail and E-commerce:** Personalized recommendations, inventory management, and customer service automation.
- **Manufacturing:** Quality control, predictive maintenance, and process optimization.

Deep Learning Foundations for GenAI

- **Neural networks and deep learning basics**, Backpropagation and optimization techniques
- Overview of popular deep learning frameworks (**TensorFlow, PyTorch**).
- Hands-on Project: Build a basic image classifier using CNNs with TensorFlow or PyTorch, applying optimization techniques learned.

Prompt Engineering and optimisation

- Strategies for crafting effective **prompts** to optimize AI responses and improve user interaction
- **Techniques for prompt optimization** and prompt chaining.

Transformers and Attention Mechanisms

- Introduction to transformers and self-attention, **BERT**, **GPT**, and other transformer-based architectures, **Hands-on exercise:** Fine-tuning a transformer model for text generation.
- **Hands-on Project:** Fine-tune a transformer model on a custom text dataset for a text classification or summarization task.

Large Language Models (LLMs)

- Overview of LLMs like **GPT-3**, **ChatGPT**, and **LLaMA**
- **Fine-tuning LLMs** for specific use cases, Applications in chatbots, summarization, and sentiment analysis.
- **Hands-on exercise:** Design and deploy a conversational AI model that simulates a customer service assistant for a chosen domain (e.g., retail or tech support).

Integration of OpenAI APIs

- Techniques for effectively integrating **OpenAI APIs**, including authentication and best practices for data handling. Rate limits and error handling.
- **Hands-on Project:** Develop a web app that uses OpenAI's API for content generation based on user inputs, including authentication and error handling.

Building Applications with LangChain

- Overview of **LangChain's** functionality for large language models.
- **Chain concepts:** sequential, memory, and conditional chains, Creating custom prompts and workflows.
- **Hands-on Project:** Personalized Financial Advice - Based on customer profiles (age, income, risk appetite), offer tailored financial advice or investment recommendations.

Utilizing Hugging Face

- Accessing and deploying **Hugging Face** pre-trained models, Fine-tuning models on custom datasets, Using Hugging Face's pipelines for **rapid deployment**, Introduction to Hugging Face's Model Hub and Transformers library.
- **Hands-on Project:** Fine-tune a Hugging Face model for a sentiment analysis task and deploy it using the Hugging Face API.

GANs (Generative Adversarial Networks)

- Fundamentals of **GANs** and their components (Generator and Discriminator).
- Variants of GANs (**DCGAN**, **StyleGAN**, **CycleGAN**).
- Applications of GANs in image synthesis, art creation, and more.
- **Hands-on exercise:** Building a simple GAN for image generation.
- **Hands-on Project:** Build and train a DCGAN to generate new images from a specific dataset, such as handwritten digits or facial images.

Variational Autoencoders (VAEs)

- Introduction to **VAEs** and their architecture
- Comparison of VAEs with GANs, Applications in anomaly
- detection and data compression.
- **Hands-on exercise:** Building a VAE for image reconstruction.

Retrieval-Augmented Generation (RAG)

- Introduction to **RAG**.
- How RAG combines generative and retrieval-based techniques, Implementing a basic RAG model for a knowledge-based task.
- **Hands-on Project:** Build a RAG-based FAQ Assistant to answer customer questions by retrieving information from a domain-specific FAQ dataset, merging retrieval and generation for clear, concise responses

Section 2: Agentic AI

Understanding Agentic AI and Autonomous Agents

- Core principles of agentic systems (Autonomy, Goal-oriented behavior, Decision-making)
- **Use cases in business and technology**

Introduction to Agentic AI

- **Task Management** and Automation
- Planning, Execution, and Optimization
- Decision-making frameworks
- **Practical demonstration** using LangChain

Key Components of Agentic AI

- **LangChain**: Introduction and hands-on exercises
- **AutoGPT**: Setup and practical application
- **BabyAGI**: Basic usage and task automation
- **CrewAI**: Introduction and practical demonstrations

Essential Tools for Agentic AI

- Integrating agents with OpenAI APIs
- Building custom agents for specific tasks
- **Hands-on session** with OpenAI's GPT models

Agentic AI Integration

- Integrating agents with **OpenAI APIs**
- Building custom agents for specific tasks
- **Hands-on session** with OpenAI's GPT models

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Build an autonomous assistant to automate content generation and social media posting using AutoGPT and LangChain.

PORTFOLIO PROJECT

DURATION 2 HRS

Develop an intelligent customer support agent that autonomously handles FAQs and escalates complex queries using BabyAGI and GPT API.

PORTFOLIO PROJECT

DURATION 1.5 HRS

Create collaborative workflows and team-based tasks automation using CrewAI.

GenAI Copilot, Deployment and Ethical Considerations in GenAI

- Learn **Model Deployment for Generative AI**, covering deployment techniques, best practices, API setup, and scaling infrastructure.
- Explore **GenAI Copilot Tools** for Coding, including GitHub Copilot, Tabnine, Amazon CodeWhisperer, Replit Ghostwriter, and Codex.
- Gain insights into **Ethical Considerations in GenAI**, addressing bias, privacy, and responsible AI development.

TERM 4 Data Visualization & Data Analysis

We will cover data visualization and data analysis including **interactive dashboards, statistical analysis, data mining, and predictive modeling** using tools like **Tableau, Power BI, Python (Matplotlib, Seaborn), and SQL.**

Module Duration:

80-85 hrs of live Instructor-led classes.

8 hrs of **on demand 1-1 session** for topics where you need help.

Number of Projects with dedicated mentor

7 practice projects

2 capstone projects on real data

Section 1 - SQL and Databases

- **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
- **Advance SQL**
 - Advance SQL Operations
 - Data Aggregations and summarizing the data
 - Ranking Functions: **Top-N Analysis**
 - Advanced SQL Queries for Analytics
- **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
- **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**
 - Mathematical Functions
 - Variables
 - Conditional Logic
 - Loops
 - **Custom Functions**
 - Grouping and Ordering
- **Programming with SQL**
 - Partitioning
 - Filtering Data
 - **Subqueries**

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Patient Medical Data Analysis with Advanced SQL

Description: Utilize advanced SQL queries to analyze patient records and identify trends in healthcare outcomes.

Tools covered: SQL (SELECT, WHERE, ORDER BY, GROUP BY, HAVING, INNER and OUTER JOINS, Aggregate Functions, Ranking Functions, Subqueries, Conditional Logic).

RESEARCH PROJECT

DURATION 2 HRS

Student Academic Performance Tracking using SQL and MongoDB

Description: Manage and analyze student data to monitor academic progress and attendance using SQL and NoSQL databases.

Tools covered: SQL (CRUD operations, JOINS, Subqueries, Variables, Conditional Logic), MongoDB (CRUD operations, Collections, Documents, JSON data handling).

Section 2 - MongoDB

• 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

• 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**

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Marketing Campaign Analytics with MongoDB

Description: Use MongoDB to store and analyze customer interactions and engagement data from marketing campaigns.

Tools covered: MongoDB (CRUD operations, aggregation framework), handling unstructured data, MapReduce functions, querying nested documents.

RESEARCH PROJECT

DURATION 2 HRS

E-commerce Product Catalog Management

Description: Design and implement a MongoDB database to manage and query a dynamic product catalog for an e-commerce platform.

Tools covered: MongoDB (CRUD operations, collections, documents, indexing), handling JSON data, aggregation pipeline.

Section 3 - PowerBI

• 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

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Music Streaming Trends Dashboard

Description: Create a dashboard to track and display music streaming statistics across different artists and genres.

Tools covered: Power BI Desktop, Data Modeling, DAX functions, Visualizations (area charts, pie charts, filters).

RESEARCH PROJECT

DURATION 2 HRS

University Enrollment Trends Dashboard

Description: Create a dashboard to track and display student enrollment trends across different faculties over time.

Tools covered: Power BI Desktop, Data Modeling, DAX functions, Visualizations (line charts, pie charts, time slicers).

Section 4 - BigData and Spark Analytics

- **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

Section 5 - Time Series

- **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**

HANDS-ON LAB PROJECT

DURATION 1.5 HRS

Time Series Forecasting of Retail Sales using Apache Spark

Description: Analyze historical retail sales data to predict future sales trends using Spark's time series capabilities.

Tools covered: Apache Spark (PySpark), Spark MLlib, Time Series

Section 6 - Tableau

- Dashboard and Stories
- Visual Analytics
- Custom Geocoding
- Polygon Maps
- WMS and Background Image
- Assignments
 - Connecting data source
 - Working with various charts
 - Deployment of Predictive model in visualization

DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

Employee Performance and Feedback Management System

Description: Design a system that stores employee performance data in SQL and feedback records in MongoDB. Use SQL queries for performance analysis and MongoDB for sentiment analysis of feedback. Present insights using a Power BI dashboard with advanced DAX calculations.

Tools covered: MySQL, MongoDB, Power BI, DAX



DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

DATA ANALYSIS

E-Commerce Data Warehouse and Reporting System

Description: Create an end-to-end data warehouse for an e-commerce company using SQL databases to store transactional data and MongoDB for customer feedback data. Develop a Power BI dashboard to visualize sales, customer trends, and feedback analysis.

Tools covered: PostgreSQL, MongoDB, Power BI, DAX



We will cover AI tools fundamentals and advanced features including **machine learning frameworks**, **natural language processing**, and **computer vision applications**.

Module Duration:

70-75 hrs of live Instructor-led classes.

10 hrs of **on demand 1-1 session** for topics where you need help.

Number of Projects with dedicated mentor

7 practice projects

2 capstone project to work on real data

Section 1 - Deep Learning using TensorFlow

- **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

- **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

- **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

- **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
- **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
- **Understanding Keras API for implementing Neural Networks**
 - Build RNN with PyTorch
 - LSTM in PyTorch
 - LSTM from CPU to GPU in PyTorch

HACKATHON PROJECT

DURATION 1.5 HRS

Image Classification with CNN

Description: Train a CNN on a dataset (e.g., CIFAR-10 or MNIST) to classify images. Utilize transfer learning with a pre-trained model like ResNet or VGG to improve accuracy.

Tools covered: TensorFlow, Keras, Python.

GROUP PROJECT

DURATION 2.5 HRS

Time Series Forecasting with LSTM

Description: Build and train an LSTM model to forecast time-series data, such as stock prices or weather trends. Evaluate the model's performance using metrics like RMSE or MAE.

Tools covered: TensorFlow, Keras, Pandas, NumPy

Get **personalized 1:1 mentorship** on **real-time practical projects** for deeper understanding and skill enhancement.

Section 2 - Natural Language Processing (NLP)

- **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
- **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
- **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

GROUP PROJECT

DURATION 2.5 HRS

Sentiment Analysis on Social Media Data

Description: Perform sentiment analysis on Twitter data using RNN or BERT to classify sentiments.

Tools covered: TensorFlow, Hugging Face, NLTK

CAPSTONE PROJECT

DURATION 2.5 HRS

GPT-3 Based Text Generation

Description: Fine-tune GPT-3 to generate domain-specific text, such as legal documents or technical reports.

Tools covered: OpenAI GPT-3, Hugging Face Transformers

- **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 Filtering**

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

- **Image Processing**

- Histogram equalization
- Thresholding and Convolution
- Sharpening and edge detection
- Morphological transformations
- Image pyramid

- **Image Classification and segmentation**

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

CAPSTONE PROJECT

DURATION 2.5 HRS

Object Detection with YOLO

Description: Implement YOLO (You Only Look Once) for real-time object detection on an image dataset. Train the model to detect multiple objects and label them with bounding boxes

Tools covered: TensorFlow, Darknet, OpenCV

Section 4 - Reinforcement Learning

- **What is Reinforcement Learning - Basics**

- Setting up Environment & Installing **OpenAI Gym**.
- OpenAI Gym Basics.
- Terminology & Environment.
- **Dynamic Programming** - Prediction, Control, and Value Approximation

- **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
- Building Blocks of Reinforcement Learning
- OpenAI Gym Tutorial Random Search
- Markov Decision Processes
- **Monte Carlo Methods**

CAPSTONE PROJECT

DURATION 2.5 HRS

Autonomous Driving Simulation with Deep Q-Network (DQN)

Description: Develop an autonomous driving agent using a Deep Q-Network to learn optimal driving strategies within a simulated environment. Train the model to avoid obstacles and maximize the reward.

Tools covered: **Python, TensorFlow/PyTorch, OpenAI Gym.**

Hands-On Experience: Gain practical skills by working on real-world assignments under expert supervision.

- **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
- **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)
- **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)
- **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)

GROUP PROJECT

DURATION 2.5 HRS

Game Environment Solver with Q-Learning

Description: Implement a Q-Learning algorithm to train an agent to solve a grid-world or maze environment using OpenAI Gym.

Tools covered: Python, OpenAI Gym, TensorFlow/PyTorch

DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

AI-Powered Customer Support Chatbot

Description: Build a chatbot using GPT-3 for query handling and RNN-based sentiment analysis. Deploy the model on AWS SageMaker with a Flask API for real-time interactions.

Tools covered: OpenAI GPT-3, TensorFlow, Flask, SageMaker, Docker



Includes AI Co-Lab real startup challenges.



DOMAIN-SPECIFIC PROJECT

DURATION 2 HRS

E-Commerce Image Classification and Sales Forecasting

Description: Create a system that classifies product images using CNN (ResNet) and forecasts sales trends using LSTM. Deploy on AWS with CI/CD pipelines.

Tools covered: TensorFlow, Flask, AWS SageMaker, Docker, Jenkins



Completing the **capstone project** helps you understand AI tools better and showcase your skills. It also strengthens your **portfolio** and supports your **career growth**.

Note: You can attain the project session in **2 modes**

1. **Live Online** with industry-led session
2. **Offline session** in 7+ cities (Instructor-Led)

We will cover Excel fundamentals and advanced features including **formulas, functions, pivot tables, and data visualization**. You'll work on several guided projects to apply these skills in practical scenarios.

Module Duration: 30-35 hrs of live Instructor-led classes.
2 practical projects with dedicated mentor

Section 1: Excel for Data Science and AI

- **Introduction to Excel for Data Science**
 - Overview of Excel Interface and Functions
 - Data Types and Data Entry
 - Basic **Formulas** and **Functions**
- **Data Cleaning and Preparation in Excel**
 - Data Importing and Exporting
 - **Handling Missing Data**
 - Data Validation Techniques
- **Data Analysis with Excel**
 - Descriptive Statistics (Mean, Median, Mode)
 - Using Excel for **Statistical Analysis**
 - Data Filtering and Sorting
- **Advanced Excel Functions**
 - Lookup Functions (**VLOOKUP**, HLOOKUP)
 - Conditional Functions (IF, SUMIF, COUNTIF)
 - Text Functions (CONCATENATE, LEFT, RIGHT)
- **Data Visualization with Excel**
 - Creating Charts and Graphs (Bar, Line, Pie)
 - Advanced Charting Techniques
 - **Conditional Formatting** for Data Insights
- **PivotTables and PivotCharts**
 - Creating and Using PivotTables

- Analyzing Data with **PivotCharts**
- Slicers and Timelines for Data Segmentation
- **Data Automation in Excel**
 - Using **Macros** for **Automation**
 - Introduction to VBA (Visual Basic for Applications)
 - Automating Repetitive Tasks
- **Project Management in Excel**
 - Creating Project Plans and Timelines
 - Tracking Progress with **Gantt Charts**
 - Resource Allocation and **Budgeting**
 - Using Excel Templates for Project Management
- **Case Studies and Applications**
 - Real-world Applications of Excel in Data Science
 - Case Studies from Various Industries
 - **Best Practices for Data Analysis in Excel**

GROUP PROJECT

DURATION 2.5 HRS

Game Environment Solver with Q-Learning

Description: Implement a Q-Learning algorithm to train an agent to solve a grid-world or maze environment using OpenAI Gym.

Tools covered: Python, OpenAI Gym, TensorFlow/PyTorch

CAPSTONE PROJECT

DURATION 2.5 HRS

Autonomous Driving Simulation with Deep Q-Network (DQN)

Description: Develop an autonomous driving agent using a Deep Q-Network to learn optimal driving strategies within a simulated environment. Train the model to avoid obstacles and maximize the reward.

Tools covered: Python, TensorFlow/PyTorch, OpenAI Gym.

We will cover data structures and algorithms using popular languages like **Python and Java**. You'll understand concepts such as **arrays, linked lists, trees, sorting, and searching algorithms**.

Module Duration:

40-50 hrs of live Instructor-led classes.

4 real time projects to practice all the concepts

Section 1: Fundamentals of Data Structures

- **Introduction to Data Structures**

- Arrays, **Linked Lists** (Singly, Doubly, Circular)
- Stacks and Queues: Implementation and Applications

- **Hands-On Lab:**

- Implement a dynamic array and stack using Python.

CAPSTONE PROJECT

DURATION 2.5 HRS

Library Management System using Stacks and Queues

Description: Build a basic library management system using stacks for book logs and queues for customer requests.

Tools covered: Python/Java

Section 2: Advanced Data Structures

- **Trees**

- Binary Trees, **Binary Search Trees**, AVL Trees, and B-Trees
- Tree Traversals (In-order, Pre-order, Post-order, Level-order)

- **Graphs**

- Graph Representations: **Adjacency Matrix**, Adjacency List
- Depth-First Search (DFS), Breadth-First Search (BFS), Shortest Path Algorithms (Dijkstra, Bellman-Ford)

- **Hands-On Lab:**

- **Implement BFS and DFS on a graph.**

CAPSTONE PROJECT

DURATION 2.5 HRS

City Navigation System using Graph Algorithms

Description: Implement a city navigation system that finds the shortest path between locations using Dijkstra's algorithm.

Tools covered: Python, NetworkX

Section 3: Sorting and Searching Algorithms

- **Sorting Algorithms**

- Bubble Sort, Merge Sort, Quick Sort, Heap Sort
- Time Complexity Analysis

- **Searching Algorithms**

- Binary Search, Linear Search, Hashing Techniques

- **Hands-On Lab:**

- Implement Quick Sort and Merge Sort, compare their time complexities.

CAPSTONE PROJECT

DURATION 2.5 HRS

E-Commerce Product Search Engine

Description: Build a product search engine that allows searching and sorting products using various algorithms.

Tools covered: Python/Java

Section 4: Dynamic Programming and Greedy Algorithms

- **Dynamic Programming**

- Introduction, Memoization, Tabulation
- Popular Problems: Longest Common Subsequence, Knapsack

- Problem, [Matrix Chain Multiplication](#)

- **Greedy Algorithms**

- Introduction and Key Concepts
- Popular Problems: Fractional Knapsack, Activity Selection

- **Hands-On Lab:**

- Solve the Knapsack problem using dynamic programming and greedy techniques.

CAPSTONE PROJECT

DURATION 2.5 HRS

Optimization Problem Solver

Description: Develop an optimization solver that uses dynamic programming and greedy algorithms for cost optimization scenarios in logistics.

Tools covered: Python/Java

We will cover project management fundamentals and methodologies like **Agile, Scrum, and Waterfall**. You'll work on several **guided projects** to apply these skills in real-world scenarios.

Module Duration:

30-35 hrs of live Instructor-led classes.

5 hrs of **on demand 1-1 session** for topics where you need help.

This project is specially designed for **mid to senior level professionals** to acquire new skills that are critical for **managing large-scale projects**

Section 1 - Project Management

Overview: Key principles of project management tailored specifically for data science projects.

Topics Covered:

- Project scope and objectives
- Timeline management
- Resource allocation and tracking
- Milestones and deliverables

Section 2: Agile and Scrum Methodologies

- **Overview:** Applying **Agile** and **Scrum** practices to manage and deliver data science projects efficiently.

- **Topics Covered:**

- Agile principles and Scrum framework
- **Sprint planning**, daily stand-ups, and retrospectives
- Managing iterative development in data science
- Role of product owners and scrum masters in data science teams

Section 3: Risk Management

Overview: Identifying potential risks in data science projects and how to mitigate them.

Topics Covered:

- Risk identification techniques
- Risk assessment and prioritization
- Mitigation strategies for data quality, model failure, and project delays

Section 4: Stakeholder Management

Overview: Effectively communicating with and managing stakeholder expectations throughout a data science project.

Topics Covered:

- Identifying key stakeholders, Managing communication and feedback loops, Aligning project goals with stakeholder expectations
- Delivering project updates and reports

GROUP PROJECT

DURATION 2.5 HRS

Mobile App Development Project using Agile Methodology

Description: Plan and manage the development of a mobile application from concept to deployment using Agile practices like Scrum.

Tools covered: Agile methodologies (Scrum), project management software (e.g., Jira, Trello), sprint planning, daily stand-ups, backlog management.

Executive-level real-time **Industrial Projects**

#1



Predictive Maintenance for Electric Vehicles

Implement a predictive maintenance system to forecast and prevent potential failures in electric vehicle components.

Tools:  

Outcome: Reduced maintenance costs and downtime, improved vehicle reliability and customer satisfaction.

#2



Supply Chain Optimization

Analyze and optimize supply chain operations to reduce costs and improve efficiency using historical sales and logistics data.

Tools:   python

Outcome: Streamlined supply chain processes, reduced operational costs, and improved product availability.

#3



Customer Purchase Prediction

Develop a machine learning model to predict customer purchase behavior based on historical data and browsing patterns

Tools:  TensorFlow  Amazon SageMaker

Outcome: Enhanced marketing strategies and personalized recommendations, leading to increased sales and customer satisfaction

#4



Drug Discovery Acceleration

Utilize AI and machine learning to accelerate the drug discovery process by predicting the efficacy of potential compounds

Tools:   IBM Watson

Outcome: Faster time-to-market for new drugs and significant cost savings in research and development.

Executive-level real-time **Industrial Projects**

#5



Employee Productivity Analysis

Analyze employee performance data to identify factors that impact productivity and develop strategies to enhance efficiency

Tools:  Power BI   python

Outcome: Improved employee performance and productivity, leading to higher overall organizational effectiveness.

#6

NETFLIX

Content Recommendation Engine Enhancement

Enhance the recommendation engine to provide more accurate and personalized content suggestions to users

Tools:  python  kafka 

Outcome: Increased user engagement and retention, leading to higher subscription rates.

#7

JPMORGAN
CHASE & CO.

Fraud Detection System

Implement an advanced fraud detection system to identify and prevent fraudulent transactions in real-time

Tools:  TensorFlow  python

Outcome: Reduced financial losses due to fraud and increased customer trust and security.

#8

Uber

Dynamic Pricing Optimization

Develop a dynamic pricing model that adjusts fares based on demand, supply, and other external factors

Tools:  MATLAB  HIVE

Outcome: Maximized revenue and improved service availability during peak times.

Executive-level real-time **Industrial Projects**

#9



Ad Spend Optimization

Develop a data-driven strategy to optimize ad spend across different channels by analyzing performance metrics and customer behavior

Tools:  Google Big Query   python

Outcome: Increased ROI on advertising campaigns and more effective allocation of marketing budgets.

#10



User Sentiment Analysis

Analyze user posts and comments to gauge public sentiment and identify trends and patterns

Tools:  python  NLTK  STORM

Outcome: Better understanding of user preferences and improved content and advertising strategies.

#11



AI-based Financial Portfolio Optimization

Develop a system that uses AI to optimize financial portfolios by analyzing market trends, risk factors, and investor preferences

Tools:  TensorFlow  python

Outcome: A web app suggesting tailored financial portfolios for CapitalOne clients based on risk tolerance and goals.

#12



AI tool for monitoring mental health

Develop an AI tool to monitor user engagement with meditation and mental health content to offer personalized well-being suggestions

Tools:  Flutter  React Native  NLP

Outcome: Tracks user interactions, sentiment, and usage patterns to offer personalized mental health advice, detect distress signs, and suggest interventions.

Thank you!

For more queries and
information please reach out
to us at:

+91 77956 87988

Visit us at

www.learnbay.co

