DATA SCIENCE AND AI CERTIFICATION PROGRAM

Master's program

Designed For Working Professionals

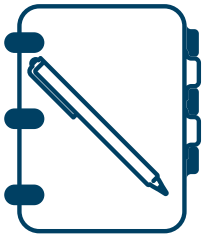
- 250+ hrs of live classroom by industry expert
- 1 year of unlimited flexible classroom subscription
- 12 + real-time projects & Capstone project
- Guaranteed job referrals in top companies

Classroom Training in Bangalore | Live Online Training | 6 Months Certification Program

Program Highlights

Learnbay offers Industry Accredited Data Science and Artificial Intelligence Certification Program which is designed for **working professionals**.Course features 12+ real world industry projects and 2 capstone projects under the mentor-ship and guidance of Data Science and AI experts.

Course is especially designed for working professionals having **1+ years of experience** in any domain. Our course is best suited for professionals looking to change their current domain and start a new career in Data science and Artificial Intelligence.



Live Sessions By Expert

- Classroom training in Bangalore
- Live Faculty led Online Training
- 250+ hrs of Interactive Classes

Project Based Learning

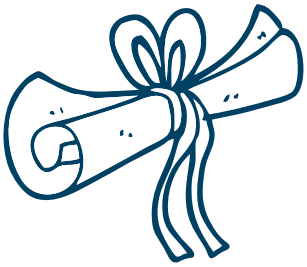
- 12+ Real World Industry Projects
- 2 Capstone Projects
- Mentorship & Guidance By Expert

One Year Flexible Subscription

- Flexibility to attend multiple batches from different trainers.
- Life time access to Recordings

Special Support to Non Programmers

- Learn Python from scratch
- Special classes for Non programming background students
- Real time Use Cases from multiple domain



Global Certification in Data Science And AI

- Certified Data science and AI program.
- Industry Accredited Global Certification Course.
- In Collaboration with IBM.

Job Assistance Program For Working Professionals

- Resume support from expert
- Interview prep session and Mock interview
- **Guaranteed job referrals** for working professionals

Top Rated Training Institute in India For Data Science And AI Certification

Become Certified Data Sience & AI Expert

4.8 ★★★★★
300+ user Review

Quora
Top Rated

4.9 ★★★★★

Click to read reviews

Program Details

Program Eligibility

Work Experience :

- Working Professionals With **1+ Years of experience** in any domain (tech or non technical)

Academics :

- BE/B.Tech (*from any branch*) , BBA/MBA, MCA/M.Tech, B.Com, Graduation in Mathematics, Statistics, IT

Who Should Apply

- Software developers/Programmers, Project Managers, Manual And Automation Test Engineer, Java and .net Developer, Informatica, Business Analyst.
- Database Admin, System Admin, Professionals from Sales, Marketing, Operations.
- SAP domain expert, Python , Embedded developer , Android/ios developer.
- Professionals from BFSI, Supply chain, Retail, healthcare, Pharma.
- Manufacturing, Mechanical, Electrical, Automobiles, Telecom domain.We have domain specific project from these sectors.
- Professionals planning for *Masters or higher education* in data science and AI

To check your eligibility, Apply for Profile Review and Counselling with expert:

Click here to apply for profile review

About Instructors

Our instructors are **working professionals** graduated from premier institutes like BITS Pilani, IIT Roorkee and working in companies as **Data Scientist/Machine Learning Engineer and Artificial Intelligence expert.**

Instructors Working in

Morgan Stanley, American Express, InMobi, Mu Sigma, Sigmoid, Samsung, HSBC

Course Prerequisite

There is **no Prerequisite** for this course as we cover programming and statistics from basics. We provide special classes & support for professionals from **non-programming/ non-technical background.**

Fees and Duration

Weekday Batches : 4 Months
Monday - Friday - 2 hours everyday

Weekend Batches : 6 Months
Saturday & Sunday - 4 hours everyday

Program Fee: Rs. 59,000/-

To know more about applicable discount, next batch details...
Live chat on Whatsapp

Click to Whatsapp

Modules And Tools

PYTHON

STATISTICS

MACHINE LEARNING

DEEP LEARNING USING TENSORFLOW

TIME SERIES ANALYSIS AND FORECASTING

NATURAL LANGUAGE PROCESSING

R

SQL

mongoDB

tableau

Apache Spark

hadoop

Power BI

github

Real world Industry Project from multiple domains

Banking

Finance

Insurance

Retail

Supply chain

Healthcare

Telecom

Manufacturing

E-commerce

Automotive

Interview Prep & Job Assistance Program

Resume Prep Session

Interview Prep Session

Mock Interviews By Expert

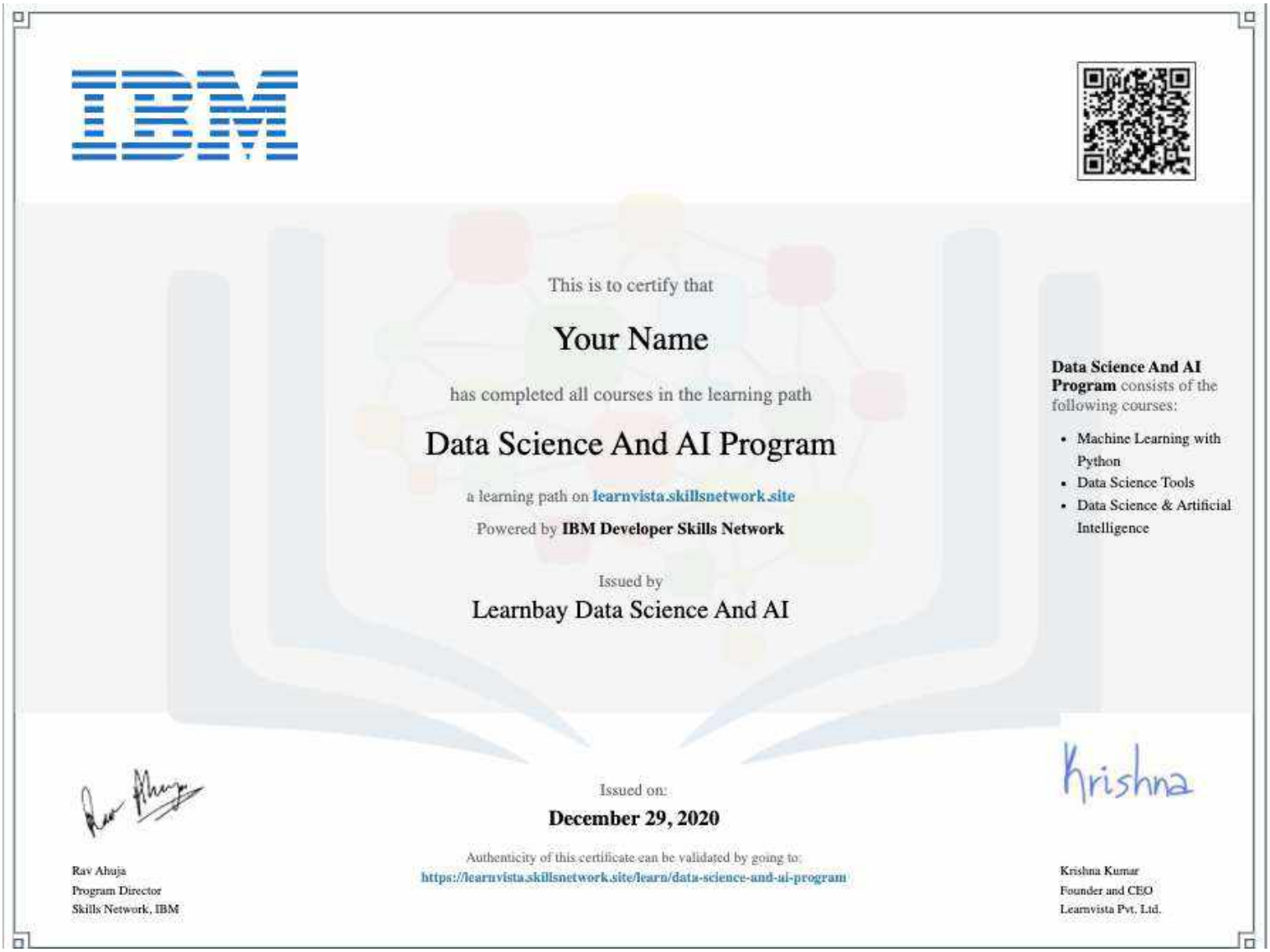
Job Referrals in data science

Global Certification in Data Science And AI

Become an industry expert with **Data Scientist & AI Master's Program** in collaboration with IBM. Upon completion of this Program, you will **receive the certificate from IBM** which will help you to become industry ready.

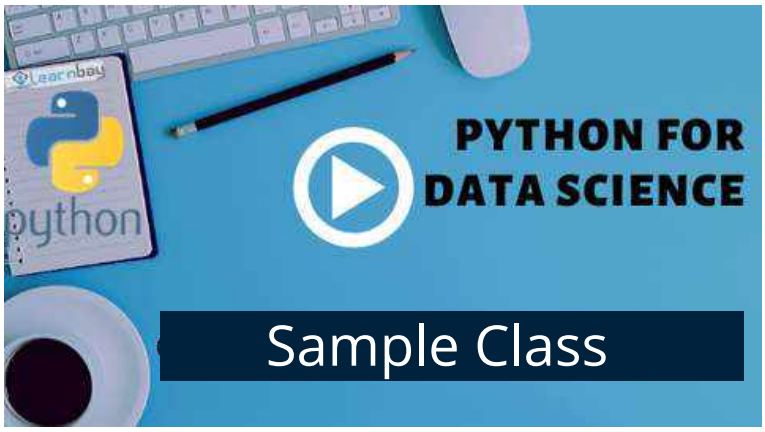
Get **Industry-renowned global certification** in Data Science and Artificial Intelligence. Our certification is **recognized globally** and industry wide in companies like JP Morgan, Morgan Stanley, Wells Fargo, Antuit , Genpact, Cognizant, Delloite, E&Y, Tredence Analytics, Mu-sigma and other **top MNCs and Banking & Finance companies**.

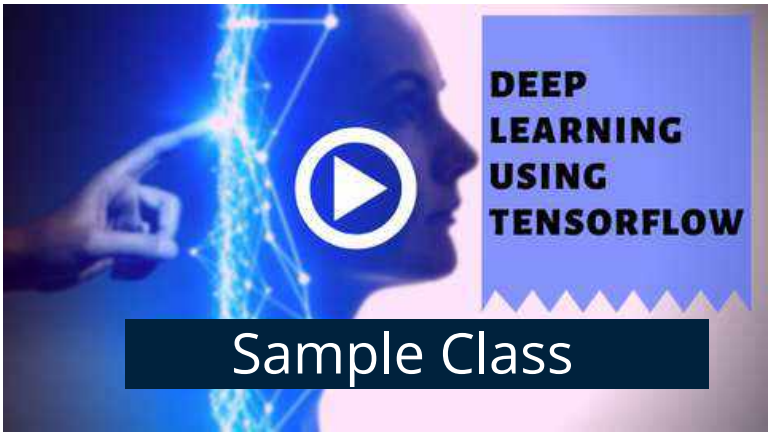
Sample Certificate



Download Certificate

Demo & Sample Class Recordings





Watch more demo session

Job Assistance

1	2	3	4	5
Certificate	Project	Resume Update	Preparation	Job Referral
After completion of your program you have to pass final exam to get IBM Certificate.	Attend project sessions from industry experts to get a hands on experience of real time projects.	After certification and project session update your resume.	Start preparing youself with mock interviews and guided interview sessions.	Once you get eligible, you will start getting guarenteed Interview Calls

Eligibility Criteria

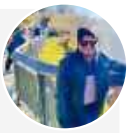
- Should have completed Term 1,2 and Term 3 of our program (Refer Course brochure for details)
- Should have more than **1 Years of work experience** (in any Domain)
- Should have scored passing marks in **final Certification exam**
- Should have completed 70% of Assignments and case studies
- At-least completed 2 Projects (Mentored and guided by our expert)

To know more about Guaranteed Interview call, Job Referral & Industrial Projects

Download Project & Job Referral Brochure

Whatsapp Now

Placement And Success Stories

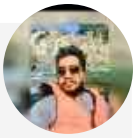


Manu Agrawal

Working at Microsoft

Everything about this program is credible
If you miss any class you can watch recorded sessions
All practice and real time codes are available in repository and the best part is you can shift batches as per your convenience.

Click here to view LinkedIn profile



Rahul Anand

Working at Affine Analytics

Learnbay is one of the best institutes in Bangalore. The faculty members are experienced working professionals and they help you to build the concepts in order to achieve your goals. The whole course and practical sessions are very helpful specially in the field of data science.

Click here to view LinkedIn profile

Pawan Yadav

Working at Oracle

I have done Data Science certification and i placed in Oracle. Journey was really tough for me because i was from core electronics domain. Mentors are really helpful and they have good knowledge. Personally i liked teaching style of Trainer Nishant. Facility of recording classes is very useful.

Click here to view LinkedIn profile

Afrin Sultana

Working at Fossil

It's a very good place to start with..LB does what it says.They have good faculties for machine learning, statistics, python and some good project sessions as well. Krishna and Abhisekh helped till I got placed. I have got multiple offers after doing the course from here and some extra effort from my end as well.So nothing is bad about it.In one word I would say it's excellent.

Click here to view LinkedIn profile

Neelesh Dugar

Working at Act21 Softwares

Very well designed and structured. I really appreciate him and would want to put some light on Utkarsh Kulshrestha. Cheers to you guys! I had an amazing experience at Learnbay, which got me where I am today. Thank you to each one of you and also Abhishek who is handling very well. All the best guys!!

Click here to view LinkedIn profile

Deevraj

Working at Mindtree

Quality of content is very nice mainly instructor concentrating of theory part, live project sessions make you feel confident to attend interviews.
Multiple batch options, access for any instructor class videos or materials. Totally positive environment around. One can join here with no second thought.

Click here to view LinkedIn profile

Keerti Bafna

Working at Antuit

I joined the Data Science batch of September 2018. The trainer was Amritansh. And since then i have evolved in Machine Learning drastically . The trainer is very educated and teaches passionately The staff is supporting and you can re-attend and switch classes anytime

Click here to view LinkedIn profile

Placement And Success Stories

Srikanth Saurav

Working at Mediamarksaturn

Machine Learning concepts & Statistics are very well explained by Utkarsh. Best thing was completing the syllabus on-time as they have promised. Trainers are clearing the doubts . Got multiple joining offers from different MNCs for Data Science and AI developer

Click here to view LinkedIn profile

Aswini Dindukurthy

Working at Deloitte

I have taken Data Science course from Learnbay 3 years back, it is Excellent training center. After my training I was equal to 3+ exp. I had a very good trainer , Real-Time Project Oriented Classes, but one thing I have to say to all that daily practice is very much needed.

Click here to view LinkedIn profile

Suman Karmakar

Working at IBM

It was a good and effective course with dedicated faculties for modules.You get flexibility to attend classes from multiple instructors.Very Supportive environment for learning.

Click here to view LinkedIn profile

Rajeev

Working at TCS

Good Trainer and nice supportive environment.One of the best classroom institute in Bangalore for working professionals looking to change their domain to data science.

Click here to view LinkedIn profile

Shakti Suwan

Working at American Express

I Joined Learnbay as Fresher and Attended training in data science And Artificial Intelligence.Course is job oriented, Practical and in-depth .To the point, well versed trainers, well engineered course. Superb!!

Click here to view LinkedIn profile

Amrita Das

Working at Cognizant

The offering here is best in the industry I would say both cost and curriculum wise. One advantage joining here is you can access their resources for lifetime unlike others where you have accessibility only for a year or so. Most importantly, there is continuous assistance for recruitment. Well,one enrolls for any course and ends up getting a handsomely paying job.

Click here to view LinkedIn profile

Read more google reviews

One Year Flexible Subscription

About One Year Classroom Subscription:

One year Flexible Subscription program is designed for working professional so that you can learn at your pace without missing any classes. With this program, you get access to attend multiple classroom/Faculty led online batches for a period of 1 year.

- Learn at your own pace with **unlimited flexible** access of multiple batches.
- Option to attend multiple batches from **different instructors** in classroom/live online mode
- Backup classes from other batches.
- You can attend **weekdays batch or weekend** or both based on your availability
- **Repeat or revise modules multiple times**.

Program Fee

Rs. 59,000 +taxes

CLICK HERE TO GENERATE DISCOUNT COUPON

PAY IN 9 MONTHS - INTEREST FREE EMI

INTEREST FREE INSTANT LOAN
WITHOUT CREDIT CARD
Aadhar Card and Pan Card required

NO COST EMI ON MAJOR CREDIT CARDS
ICICI, HDFC, RBL, Standard Chartered, Axis bank,Kotak credit cards

Click here to apply for INTEREST FREE LOAN

How to Apply For this Program ?

Talk to Our Admission Executive

Contact our Admission Team for more details on course eligibility, Queries on course curriculum, Certification etc.. If your profile is suitable for this course, you will be further guided for detailed counselling and Profile Review sessions.

Request A Callback

Whatsapp Now

Apply For Profile Review & Personalized Counselling

Attend Personalised Career Counselling and profile review session with expert. This session will help you to understand whether your profile is suitable for Data Science and AI certification course.
Note: You can attend this session online or visting our HSR center (Bangalore)

Apply For Profile Review

Pay and Enrol For this Program

Contact our Admission Officer for discount coupon. Apply the discount coupon and enroll for IBM certified Program.

Pay and Enroll for the program

Syllabus | 4 Terms | 6 Months

MODULE 0 :

Modules/Tools : **Basic programming fundamentals for Non-programming background aspirants**

Term Duration : 4 days (10 hours)

TERM 1 :

Modules/Tools : **Core Python + Numpy + Pandas + Matplotlib + Seaborn**

Term Duration : 5 Weeks (40 hours) : : 1.5 Months

TERM 2 :

Modules/Tools : **Statistics (3 weeks - 24 hrs) + Machine Learning (6 Week - 48 hrs) + Capstone Project**

Term Duration : 9 Weeks (72 hours) : : 2 Months

TERM 3 :

Modules/Tools : **Deep Learning using Tensor-flow (2 Weeks - 16 hours) + Natural Language Processing & Text Analytics (3 Weeks - 20 hours) + Capstone Project**

Term Duration : 5 Weeks (40 hours) : : 1 Month

Final Exam for Certification after Term 3

Important Note :

After Successful completion of term 1, term 2 and term 3, Candidates become eligible for Job Assistance Program (2- 3 weeks) which includes :

- Resume Session and Assistance
- Interview Prep Session & Mock Interview
- Participating in Live Kaggle Competitions
- List of Important Interview Questions from each modules
- **Guaranteed Job Referrals for Data Science/ML engineer roles**
- You can start attending interviews after Term 3 and keep learning other modules from Term 4 simultaneously.
- Attend guided session for real time projects from multiple domain and get project Support/Mentorship from expert instructors.

TERM 4 :

Modules/Tools : **(SQL + MongoDB) + (Tableau + PowerBI) + Cloud Deployment of ML Model using GCP +(Hadoop basics & Apache Spark) + R Programming**

Term Duration : 9 Weeks (72 hours) : : 2 Months

MODULE 0 : SPECIAL PROGRAMMING CLASSES |
10 hours

Module 0
Basic

Chapter 1: Introduction to Programming (3 hrs)

What is a programming language ?
Source code Vs bytecode Vs machine code
Compiler Vs Interpreter
C/C++, Java Vs Python

Chapter 2: Jupyter notebook basics (1 hrs)

Different type of code editors in pythonIntroduction to Anaconda and jupyter notebookFlavours of python.

Chapter 3: Python Programming Basics (2 hrs)

Variable Vs identifiers Vs strings
Operators Vs operand Procedure oriented Vs modular programming

Chapter 4: Statistics basics (2 hrs)

Introduction to statisticsMean, median, mode, Standard deviation, AverageIntroduction to probability, permutations and combinationsIntroduction to linear Algebra

Chapter 5: Git and GitHub (2 hrs)

Learn the key concepts of the Git source control system
Step through the entire basic Git workflow
Configure SSH for authentication
Create and use a remote repository on GitHub
Git Overview
Set up & configuration
Working with git locally

[NOTE]

This module 0 is for those who are from non-technical background like Mechanical, BBA, MBA, B.Com, M.Com, etc.
Or for those who work in Non-IT sectors to get in-depth knowledge of programming and how to use it in Data Science.

MODULE 1 : PYTHON FOR DATA SCIENCE | 40 hours

Term 1
Python

1. 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.

2. 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.

3. Strings, Decisions And 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.
Introduction to for Loops,Using continue and break.

Class hands-on :

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

4. 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

Class hands-on :

- *Program to convert tuple to dictionary*
- *Remove Duplicate from Lists*
- *Python program to reverse a tuple*
- *Program to add all elements in list.*
- *+ 3 more programs to be covered in class*

5. Functions And Modules

Introduction To Functions – Why
Defining Functions
Calling Functions
Functions With Multiple Arguments.
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 from functions, Lambda, modules, Generators and Packages.

6. File I/O And 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,re module

Class hands-on :

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

MODULE 1 : PYTHON FOR DATA SCIENCE | 32 hours

7. Data Analysis Using Numpy And Pandas

Introduction to **Numpy**. Array Creation,Printing Arrays, Basic Operation - Indexing,Slicing and Iterating, Shape Manipulation - Changing shape,stacking and spliting of array
Vector stacking, Broadcasting with Numpy, *Numpy for Statistical Operation.*
Pandas : Introduction to Pandas
Importing data into Python
Pandas Data Frames,Indexing Data Frames ,Basic Operations With Data frame,Renaming Columns,Subletting and filtering a data frame.

8. Data Visualisation using Python: Matplotlib and Seaborn

Matplotlib:
Introduction,plot(),Controlling Line Properties,Subplot with Functional Method, MUltiple Plot, Working with Multiple Figures,Histograms
Seaborn :
Intro to Seaborn And Visualizing statistical relationships , Import and Prepare data .Plotting with categorical data and Visualizing linear relationships
Seaborn Exercise

Real time Use cases in Python to be Covered in Class

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)

Assignment 1 (Week 1):

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

Assignment 2 (Week 2):

10 Python Programs and practice set on List,Tuples ,Dictionaries & matrices operations

Assignment 3 (Week 3):

10 Coding exercises on Functions, File And Regular Expression

Assignment 4 (Week 4):

15 Programs and Practice set Questions on Numpy and Pandas

Assignment 5 (Week 5):

2 Case Studies using Numpy Pandas and Matplotlib.

1. Fundamentals of Math and Probability

Basic understanding of linear algebra, Matrics, vectors
Addition and Multimplication of matrics
Fundamentals of Probability
Probability distributed function and cumulative distributed function.

Class Hand-on

Problem solving using R for vector manupulation
Problem solving for probability assignments

2. Descriptive Statistics

Describe or sumarise a set of data
Measure of central tendency and measure of dispersion.
The mean,median,mode, curtosis and skewness
Computing Standard deviation and Variance.
Types of distribution.

Class Handson:

5 Point summary BoxPlot
Histogram and Bar Chart
Exploratory analytics R Methods

3. Inferential Statistics

What is inferential statistics
Different types of Sampling techniques
Central Limit Theorem
Point estimate and Interval estimate
Creating confidence interval for population parameter
Characteristics of Z-distribution and T-Distribution
Basics of Hypothesis Testing
Type of test and rejection region
Type of errors in Hypothesis resting, conti..

conti..
Type-I error and Type-II errors
P-Value and Z-Score Method
T-Test, Analysis of variance(ANOVA) and Analysis of Co variance(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

4. Hypothesis Testing

Hypothesis Testing
Basics of 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, Effect Size, and Confidence Intervals

5. Data Processing & Exploratory Data Analysis

Introduction to Data Cleaning
Data Pre-processing
What is Data Wrangling?
How to Restructure the data?
What is Data Integration?
Data Transformation
EDA : Finding and Dealing with Missing Values.What are Outliers? Using Z-scores to Find *Outliers*. Introduction to Bivariate Analysis,Scatter Plots and Heatmaps. Introduction to Multivariate Analysis

Introduction To Machine Learning

What is Machine Learning?
Introduction to Supervised and Unsupervised Learning
Introduction to SKLEARN (Classification, Regression, Clustering, Dimensionality reduction, Model selection, Preprocessing)
What is Reinforcement Learning?
Machine Learning applications
Difference between Machine Learning and Deep Learning

1. Supervised Learning

Support Vector Machines
Linear regression
Logistic regression
Naive Bayes
Linear discriminant analysis
Decision tree
k-nearest neighbor algorithm
Neural Networks (Multilayer perceptron)
Similarity learning

2. Linear Regression

Introduction to Linear Regression
Linear Regression with Multiple Variables
Disadvantage of Linear Models
Interpretation of Model Outputs
Understanding Covariance and Colinearity
Understanding Heteroscedasticity

Case Study – Application of Linear Regression for Housing Price Prediction

3. Logistic Regression

Introduction to Logistic Regression.– Why Logistic Regression .
Introduce the notion of classification
Cost function for logistic regression
Application of logistic regression to multi-class classification.
Confusion Matrix, Odd's Ratio And ROC Curve
Advantages And Disadvantages of Logistic Regression.
Case Study: To classify an *email as spam or not spam using logistic Regression.*

4. Decision Trees

Decision Tree – data set
How to build decision tree?
Understanding Kart Model
Classification Rules- Overfitting Problem
Stopping Criteria And Pruning
How to Find final size of Trees?
Model A decision Tree.
Naive Bayes
Random Forests and Support Vector Machines
Interpretation of Model Outputs

Case Study:

1 Business Case Study for Kart Model
2 Business Case Study for Random Forest
3 Business Case Study for SVM

5. Unsupervised Learning

Hierarchical Clustering
k-Means algorithm for clustering – groupings of unlabeled data points.
Principal Component Analysis(PCA)-Data
Independent components analysis(ICA)
Anomaly Detection
Recommender System-collaborative filtering algorithm
Case Study– Recommendation Engine for e-commerce/retail chain

6. Natural language Processing

Introduction to natural Language Processing(NLP).
Word Frequency Algorithms for NLP
Sentiment Analysis
Case Study :
Twitter data analysis using NLP

7. 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

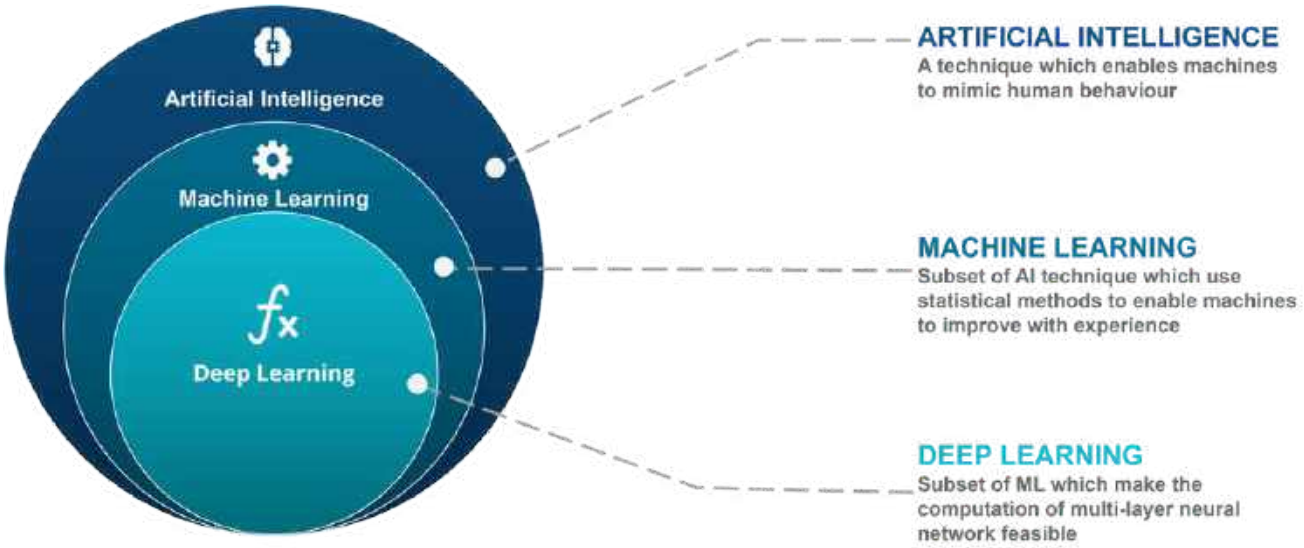
8. ARIMA and Multivariate Time Series Analysis

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 : Performing Time Series Analysis on Stock Prices

Important 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.**

Assignments :
Statistics Assignments : Total 4 practice set and Assignments from Statistics
Machine Learning Assignments : Total 3 Practice Set And 2 Real time use case as Assignments

Assessment Test For Term2 :
Duration : 3 hours
Question Type : Objective & ML Case Studies



1. Introduction to Deep Learning And Tensor Flow

Neural Network
Understaing Neural Network Model
Installing TensorFlow
Simple Computation ,Contants And Variables
Types of file formats in TensorFlow
Creatting A Graph – Graph Visualization
Creating a Model – Logistic Regression
Model Building using tensor flow
TensorFlow Classification Examples

2. Introduction to Tensor Flow

Installing TensorFlow
Simple Computation ,Contants And Variables
Types of file formats in TensorFlow
Creatting A Graph - Graph Visualization
Creating a Model - Logistic Regression
Model Building
TensorFlow Classification Examples

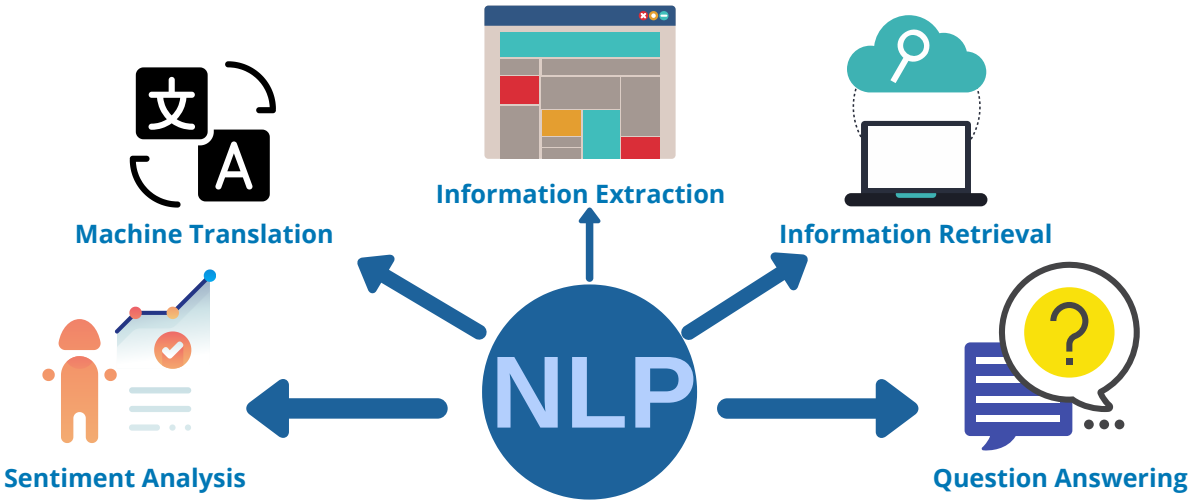
3.. Understanding Neural Networks With Tensor Flow

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 on backpropagation

4. Convolutional Neural Network(CNN)

Convolutional Layer Motivation
Convolutional Layer Application
Architecture of a CNN
Pooling Layer Application
Deep CNN
Understanding and Visualizing a CNN

Project : Building a CNN for Image Classification



1. Introduction to NLP & Text Analytics

Introduction to Text Analytics
Introduction to NLP
What is Natural Language Processing?
What Can Developers Use NLP Algorithms For?
NLP Libraries
Need of Textual Analytics
Applications of Natural Language Procession
Word Frequency Algorithms for NLP
Sentiment Analysis

2. Text Pre Processing Techniques

Need of Pre-Processing
Various methods to Process the Text data
Tokenization ,Challenges in Tokenization
Stopping ,Stop Word Removal
Stemming - Errors in Stemming
Types of Stemming Algorithms - Table
lookup Approach ,N-Gram Stemmers

3. Distance Algorithms used in Text Analytics

String Similarity
Cosine Similarity Mechanishm - Similarity
between Two text documents
Levenshtein distance - measuring the difference between two sequences
Applications of Levenshtein distance
LCS(Longest Common Sequence) Problems
and solutions ,LCS Algorithms

4. Information Retrieval Systems

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

5. Projects And Case Studies

a. Sentiment analysis for twitter, web articles
b. Movie Review Prediction
c. Summarization of Restaurant Reviews

1. RDBMS And SQL Operations :

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

2. NoSQL Databases :

Topics - What is HBase?
HBase Architecture, HBase Components,
Storage Model of HBase,
HBase vs RDBMS
Introduction to Mongo DB, CRUD
Advantages of MongoDB over RDBMS
Use cases

3. Programming with SQL :

Mathematical Functions
Variables
Conditional Logic
Loops
Custom Functions
Grouping and Ordering
Partitioning
Filtering Data
Subqueries

4. MongoDB Overview :

Where MongoDB is used?
MongoDB Structures
MongoDB Shell vs MongoDB Server
Data Formats in MongoDB
MongoDB Aggregation Framework
Aggregating Documents
What are MongoDB Drivers?

5. 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

6. Introduction to MongoDB :

What is MongoDB?
Charateristics and Features
MongoDB Ecosystem
Installation process
Connecting to MongoDB database
Introduction to NoSQL
Introduction of MongoDB module
What are ObjectIds in MongoDb

MODULE 7 : TABLEAU AND POWER BI | 16 hours

Term 4

Value Added Skillset

1. Introduction to Tableau :
Connecting to data source
Creating dashboard pages
How to create calculated columns
Different charts
Hands-on :
Hands on on connecting data source and data cleansing
Hands on various charts

2. Visual Analytics :
Getting Started With Visual Analytics
Sorting and grouping
Working with sets, set action
Filters: Ways to filter, Interactive Filters
Forecasting and Clustering
Hands-on :
Hands on deployment of Predictive model in visualization

3. Dashboard and Stories :
Working in Views with Dashboards and Stories
Working with Sheets
Fitting Sheets
Legends and Quick Filters
Tiled and Floating Layout
Floating Objects

4. Mapping :
Coordinate points
Plotting Latitude and Longitude
Custom Geocoding
Polygon Maps
WMS and Background Image

5. 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

6. Programming with Power BI :
Working with Timeseries
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

MODULE 8 : BIG DATA AND SPARK ANALYTICS | 12 hours

1. Introduction To Hadoop :
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

2. Apache Spark Analytics :
What is Spark
Introduction to Spark RDD
Introduction to Spark SQL and Dataframes
Using R-Spark for machine learning
Hands-on:
installation and configuration of Spark

Using R-Spark for machine learning programming

3. Apache Spark Analytics :
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

Hands-on:
Spark RDD programming

Hands-on:
Spark SQL and Dataframe programming

MODULE 9 : R PROGRAMMING | 12 hours

1. Introduction To R :
Installation Setup
Quick guide to RStudio User Interface
RStudio's GUI3
Changing the appearance in RStudio
Installing packages in R and using the library
Development Environment Overview
Introduction to R basics
Building blocks of R
Core programming principles
Fundamentals of R

2. Programming with R :
Creating an object
Data types in R
Coercion rules in R
Functions and arguments
Matrices
Data Frame
Data Inputs and Outputs with R
Vectors and Vector operation
Advanced Visualization
Using the script vs. using the console

3. Manipulating Data :
Data transformation with R - the Dplyr package - Part
Data transformation with R - the Dplyr package - Part
Sampling data with the Dplyr package
Using the pipe operator in R
Tidying data in R - gather() and separate()
Tidying data in R - unite() and spread()

4. Visualizing Data :
Intro to data visualization
Introduction to ggplot2
Building a histogram with ggplot2
Building a bar chart with ggplot2
Building a box and whiskers plot with ggplot2
Building a scatterplot with ggplot2

MODULE 10 : TRAINING AND DEPLOYING MACHINE LEARNING MODEL USING GCP | 8 hours

1. Introduction To GCP Cloud ML Engine :
Introduction to Google CloudML Engine
CloudML Engine in Machine Learning WorkFlow
Components of Cloud ML Engine - Google Cloud Platform Console. gcloud command-line tool and Rest API

2. Training Machine Learning Model :
Developing a training application
Packaging a training application
Running and monitoring a training job
Using hyperparameter tuning
Using GPUs for training models in the cloud

Real Time Industry Projects

1

Domain - Banking & Finance

DataSet : Banking Data

Project : Loan Default Prediction

The bank wants to improve their services by finding interesting groups of clients. Fortunately, the bank stores data about their clients, the accounts (transactions within several months), the loans already granted, the credit cards issued. This process of loan default prediction can be done with machine learning algorithms.

2

Domain - Retail industry

DataSet : BigBazar/Future Group

Project : Clustering Customers

Big Bazaar has retail outlets across major metropolitan cities in India. With the help of machine learning algorithms we can better understand customer behaviour and understand their buying needs better.
BigBazaar runs various loyalty programs, festive offers which provide their customer more opportunities to avail discounts.

BIG BAZAAR

3

Domain - Demand/Supply

DataSet : IBM

Project - IBM HR Analytics

Applying analytic processes to the human resource department of an organization in the hope of improving employee performance and therefore getting a better return on investment.
This is especially concerning if your business is customer facing, as customers often prefer to interact with familiar people.

IBM

4

Domain - Demand/Supply

DataSet : Uber & Rapido

Project- Forecasting Uber Demand

The goal is to create an interactive dashboard using Tableau
This Tableau Dashboard can be used to get historical insights into a neighborhood,
For example,
see its upcoming forecasted demand,
increase the accuracy,
decrease surge pricing events.

rapido Uber

5

Domain - Healthcare

DataSet : Samsung

Project - Analyzing Health Data and tracking human activity

The goal is to breakdown all the data that the Samsung Health app has collected and see what useful insights we can gain by analyzing it.

SAMSUNG

6

Domain - Banking & Finance

DataSet : Banking Dataset

Project - Identify fraudulent credit card transactions.

To recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase.
It involves various processes like Data Cleaning, Data Visualization, Insights generation, Model generation, Feature Engineering and so on.

7

Domain - E-Commerce
DataSet : Amazon Data

Project - Consumer Reviews of Amazon Products

The goal is to analyze Amazon's most successful consumer electronics product launches; discover insights into consumer reviews and assist with machine learning models.
What are the most reviewed Amazon products?
How do the reviews in the first 90 days after a product launch?

amazon

8

Domain - Travel & Hospitality
DataSet : Airbnb

Project - Airbnb New User Bookings

The goal is to predict which country a new user's first booking destination will be.
By accurately predicting where a new user will book their first travel experience, Airbnb can share more personalized content with their community, decrease the average time to first booking, and better forecast demand.

airbnb

9

Domain - Media and Entertainment
DataSet : Netflix

Project - Netflix Movies and TV Shows

Explore what all other insights can be obtained from the list of tv shows and movies available on Netflix as of 2019. Understanding what content is available in different countries
Identifying similar content by matching text-based features
Network analysis of Actors / Directors and find interesting insights.

NETFLIX

10

Domain - Retail
DataSet : Walmart

Project - Walmart Sales Forecasting

This dataset contains the sales for each department from the Walmart dataset containing data of 45 Walmart stores, selected holiday markdown events are also included . These markdowns are known to affect sales, but it is challenging to predict which departments are affected and the extent of the impact.

Walmart

11

Domain - Automation
DataSet : BMW dataset

Project -BMW Pricing Challenge

To find a good statistical model to describe the value of a used car depending on the basic description
How does the estimated value of a car change over time? Can you detect any patterns?
How big is the influence of the factors not represented in the data on the price?

12

Domain - Manufacturing
DataSet : Bosch

Project - Bosch Production Line Performance

To predict internal failures using thousands of measurements and tests made for each component along the assembly line. This would enable Bosch to bring quality products at lower costs to the end user.
The goal is to predict which parts will fail quality control

BOSCH

13

Domain - Social Media

DataSet : youtube

Project - Trending YouTube Video Statistics

The dataset of this project are daily record of the top trending YouTube videos, to generate insights like :
Sentiment analysis in a variety of forms
Categorising YouTube videos based on their comments and statistics
Training ML algorithms like RNNs to generate their own YouTube comments.

YouTube

14

Domain - Telecom

DataSet : Telecom

Project - Identify And Predict Customer churn in telecom industry

The goal is to develop a churn prediction model which assists telecom operators to predict customers who are most likely subject to churn.Also to understand the customer behavior and reasons for churn.Apply multiple classification models to predict the customer churn in telecom industry.

15

Domain - Supply Chain

DataSet : Dataco

Project - Smart Supply Chain for Big Data Analysis

A DataSet of Supply Chains used by the company DataCo Global is used for the analysis. Dataset of Supply Chain , which allows the use of Machine Learning Algorithms and R Software.
It also allows the correlation of Structured Data with Unstructured Data for knowledge generation.

DataCo

Contact Us

Call Us
7349-2222-63

Mail Us
contacts@learnbay.co

Visit Us
www.learnbay.co

Whatsapp Us
7349-2222-63

click on icon to
Follow Us On

Marathahalli Office :
Learnbay,19/1,2nd Floor, Classic Aura(Beside Aricent),Marathahalli - Outer Ring Road,Kadubeesanahalli, Bengaluru, Karnataka

HSR Office :
Learnbay ,147, 5th Main Rd, Rajiv Gandhi Nagar, HSR Sector 7,Near Salarpuria Serenity, Bengaluru, Karnataka 560102
INDIA
+917349222263