

# PYTHON PROGRAMMING

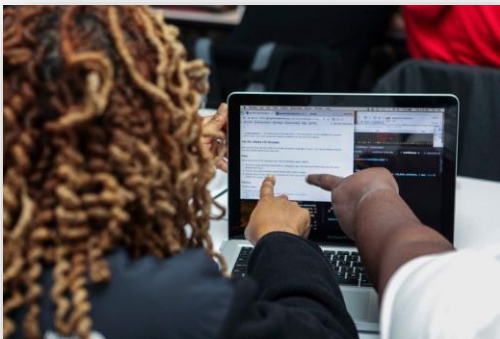
- 01 FOUNDATIONS
- 02 PYTHON FOR MACHINE LEARNING
- 03 PYTHON FOR DEEP LEARNING

# Overview

---

Learning **Python** is essential for those who wish to learn business analytics and data science as it empowers analysts and data scientists to efficiently tackle complex data challenges and uncover valuable insights that drive business growth and informed decision-making.

Python's simplicity, versatility, and the wealth of libraries and resources available make it an indispensable tool for anyone seeking to excel in data science and business analytics.



# Course 1: Foundations of Python

**Duration: 12 Hours**

This course is designed to provide participants with a solid foundation in Python programming for data science and business analytics. The course will cover essential Python concepts and foundational libraries, enabling participants to effectively manipulate, analyze, visualize data. The course will be **hands-on and project-based**, allowing participants to apply their skills to real-world data sets.

The objective of the course is to enable the participants to

1. Understand and learn core features of python.
2. Learn how to read and write various data formats.
3. Learn how to prepare and clean data using python.
4. Learn how to do basic statistical analysis to find key insights from the data using python.
5. Learn how to create basic charts and plots for univariate and bivariate analysis.

## Software:

Install latest Anaconda Distribution for Python (3.8) on your desktop / laptop

<https://www.anaconda.com/download/>

Create a working directory on your desktop/laptop, where you can store all data and the programs of this training

<b>Sessions 1&amp;2 3 Hrs</b>	Python Overview	<ul style="list-style-type: none"> <li>• Overview of python libraries</li> <li>• Core language features</li> <li>• Variables and loops</li> <li>• Functions</li> </ul>	Jupyter notebook and Google Colab; Python functions & iterators
<b>Sessions 3&amp;4 3 Hrs</b>	Advance Language Features and Introduction to Data frames	<ul style="list-style-type: none"> <li>• Writing utility functions</li> <li>• Collections: Lists, dictionaries, tuples and sets</li> <li>• Reading and Writing csv, excel files</li> </ul>	Collections and Extracting and loading stock trading data
<b>Sessions 5&amp;6 3 Hrs</b>	Advanced Data Frame operations	<ul style="list-style-type: none"> <li>• Dealing with real world data</li> <li>• Filtering, grouping, joining, sorting etc.</li> <li>• Basic statistical analysis</li> </ul>	Value at Risk Analysis for Stock Prices
<b>Sessions 7&amp;8 3 Hrs</b>	Exploratory Data Analysis (EDA) using Python	<ul style="list-style-type: none"> <li>• Univariate and Bivariate Analysis</li> <li>• Histogram, Box plot, Bar Plot, Scatter Plot, Heatmap etc.</li> </ul>	<b>Insurance Data:</b> Exploratory Analysis

# Course 2: Python for Machine Learning

**Duration: 24 Hours**

AI and Machine Learning is emerging as a hot new profession and academic discipline. Harvard Business Review says Data Scientist is the Sexiest Job of the 21st Century. But demand for prompt engineers and data scientists are racing ahead of supply. People with the necessary skills are scarce, primarily because the discipline is so advent.

## Pre-requisites:

A basic understanding of data and programming is required  
Programming knowledge using Python is essential

## Learning Goals:

- Understand the Data Science landscape and problem domain
- Access, prepare and explore data with Pandas & Scipy
- EDA: Univariate, Bivariate and Multivariate Analysis
- Feature engineering, missing values imputation, scaling and categorical encoding
- Regression, classification algorithms
- Linear Regression, KNN, Decision Tree, Logistic Regression, Random Forest
- Feature Engineering
- Clustering and Dimensionality Reduction
- Building and validating models
- Evaluation metrics
- Model Deployment

<p>Machine Learning Overview</p>	<p>Introduction to Machine Learning          Setting up Python Environment for Data Analysis          Overview of Python Stack of Machine Learning - Numpy, Pandas</p>
<p>Exploratory Data Analysis</p>	<p>Basic Statistical analysis using scipy.stats          Drawing Histograms, Bar charts, Density Plots, Box Plots          Drawing Density plots and understating data distributions          Univariate Analysis, Bivariate and Multivariate Analysis</p>
<p>Supervised Learning: Regression</p>	<p>Feature Engineering: Scaling, Categorical Encoding, Imputation          Supervised Learning: Regression          Regression: Linear Regression, KNN          Model Evaluation: K-Fold Cross Validation          Metrics: RMSE, R2 score          Creating Model Pipeline          All Hands on using Scikit-learn library</p>
<p>Supervised Learning: Classification</p>	<p>Supervised Learning: Classification          KNN, Decision Trees,          Ensemble Methods: Random Forest, Boosting          Model Evaluation: Confusion Matrix, Precision, Recall</p>
<p>Unsupervised Learning: Clustering and Dimensionality Reduction</p>	<p>Clustering: K-Means          Finding optimal Clusters: Elbow Method, Cluster Map          Dimensionality Reduction: PCA (Principal Component Analysis)</p>
<p>Model Validation and Deployment</p>	<p>Cross validations          Hyperparameter Tuning          Model Deployment</p>

# Course 3: Python for Deep Learning

**Duration: 26 Hours + Optional 5 Hours on PyTorch**

Python is essential for deep learning due to its readability and extensive libraries like NumPy, pandas, and matplotlib. Its simplicity and large community make it ideal for rapid prototyping. PyTorch, a popular deep learning framework, builds upon Python's strengths, providing dynamic computation graphs and automatic differentiation. This allows for easy experimentation and efficient training of complex neural networks. PyTorch also offers GPU acceleration, enhancing performance for deep learning tasks. Together, Python and PyTorch create a user-friendly and powerful environment for researchers and practitioners, making them indispensable tools for learning and applying deep learning techniques in various domains.

This comprehensive course is designed to equip participants with the essential knowledge and practical skills needed to excel in deep learning using PyTorch. Starting with an introduction to tensors and PyTorch basics, participants will learn to build input pipelines, design neural networks, and grasp computation graphs for efficient training. The course will cover automatic differentiation, enabling the computation of gradients for optimization.

Furthermore, participants will delve into training deep neural networks, fine-tuning hyperparameters, and applying these techniques to computer vision tasks using CNNs. Additionally, the course will explore processing sequences with RNNs and CNNs and delve into natural language processing using RNNs and attention mechanisms and transformer architecture. Throughout the course, hands-on projects and real-world applications will solidify participants' understanding and ensure they are well-prepared to tackle complex deep learning challenges.

The objective of the course is to enable the participants to:

- Understand the fundamentals of tensors and their role in PyTorch for efficient data representation.
- Master the construction of input pipelines to handle diverse datasets and streamline data pre-processing.
- Learn to design, build, and train neural networks, acquiring skills to tackle various real-world problems.
- Grasp the concept of computation graphs and their importance in optimizing deep learning models.
- Explore automatic differentiation to compute gradients, enabling efficient optimization during training.
- Gain expertise in training deep neural networks and fine-tuning hyperparameters for improved performance.
- Apply convolutional neural networks (CNNs) to solve computer vision tasks, enhancing image recognition and analysis.
- Develop proficiency in processing sequential data using recurrent neural networks (RNNs) and CNNs.
- Apply RNNs with attention mechanisms and transformers to tackle natural language processing tasks, such as language translation, sentiment analysis, Named Entity Recognition etc.
- Complete hands-on projects and real-world applications to reinforce understanding and practical application of deep learning with PyTorch.

### **Pre-requisites:**

Programming knowledge of Python and in-depth knowledge of basic machine learning concepts like supervised vs unsupervised learning, basic classifiers, and machine learning process.

### **Hardware, Software Requirement and Installation:**

A laptop running 64-bit OS (Linux/OSX/Windows)

We use Google Colab service and hence installation is not required



# What You'll Learn

## Unit 1 1.5 Hrs

### PyTorch and its programming Model

- Concept of tensors & creating it
  - Manipulating the data type and shape of a tensor
  - Applying mathematical operations to tensors
  - Split, stack, and concatenate tensors
- 

## Unit 2 1.5 Hrs

### Building input pipelines in PyTorch

- PyTorch DataLoader
  - Shuffle, batch, and repeat
  - Creating and fetching datasets
- 

## Unit 3 2 Hrs

### PyTorch's computation graphs

- Understanding computation graphs and creating one
  - Storing and updating model parameters
  - Computing gradients via automatic differentiation
  - Different modules and custom layers in PyTorch
  - Implementations of common architectures
- 

## Unit 4 2 Hrs

### Concepts of Deep Learning and Artificial Intelligence

- Introduction to Deep Learning
- Representation Learning
- Feature Engineering Vs. Automated Learning
- Deep Learning Applications and Frameworks
- Deep Learning Challenges
- Hands-On



---

**Unit 5**  
8 Hrs

Training Deep Learning & Hyperparameter Tuning

- Multi-Layer Perceptron
- Gradient Descent and Back-Propagation
- Activation Functions at Hidden and Output Layer
- Hyper-parameters Tuning: Number of Hidden Layers, Number of Neurons per Hidden Layer, Learning Rate, Batch Size, and Other Hyperparameters
- The Vanishing/Exploding Gradients Problems
- Reusing Pretrained Layers
- Faster Optimizers
- Learning Rate Scheduling
- Avoiding Overfitting Through Regularization - L1, L2, Dropout, Batch Normalization
- Unsupervised Learning using AutoEncoders and Transfer Learning
- Applications and Hands-On

---

**Unit 6**  
8 Hrs

Deep Convolutional Neural Network

- Convolutions - Filters, Feature Maps, Padding, Striding
- Convolution over multiple channels
- Convolution Neural Network - The building blocks of CNNs
- Subsampling layers (Pooling Layers)
- Activation Functions
- Flattening and Fully Connected Layer
- Receptive Field
- Significance of multiple layers
- MLP Vs. CNN -- Issue with FC layer and CNN with only convolution layers
- 1 \* 1 Convolution
- Fully Convolution Layers and Global Average Pooling
- CNN Architectures - VGG, Inception, ResNet, Xception
- Applications and Hands-On

---

**Unit 7**  
8 Hrs

Processing and Modelling Sequential Data using Convolutional Neural Network, Recurrent Neural Networks and Transformers

- Natural Language Processing
- Pipelines for Text Data
- Types of encodings and its shortcomings
- Word Embeddings (Static): Word2Vec/Glove
- Recurrent Neurons, Layers and Recurrent Neural Networks (RNNs)
- Training and Types of RNNs
- RNN Architectures - (LSTM)/ (GRU)
- Stacked RNNs, Bidirectional RNNs, Stateful RNNs
- Encoder-Decoder Network
- Teacher Forcing
- Beam Search - Avoiding Greedy Approach
- Model Evaluation - BLEU score
- Attention Mechanism - Bahdanau and Luong Attention
- Transformer Models
- Autoencoding and Auto-regressive Family Models
- Applications and Hands-On using HuggingFace library

## About Instructors



### Manaranjan Pradhan

Manaranjan Pradhan, an IIM Bangalore alumnus, has about 20+ years of industry experience working on **Big Data & Machine Learning**. He has worked with TCS, HP, and iGATE and worked on large scale project implementations for customers like Motorola, Home Depot, CKWB Bank, P&G in the roles of solution and technical architect. He is a freelancer who provides consulting and training on big data & Data Science including Machine Learning. He has been teaching Big Data and Machine Learning for 7+ years and has **trained more than 1000 people** from several large MNCs including *EMC, CISCO, TESCO, HP, YODLEE, Goldman Sachs, Software AG, Amadeus, Cognizant, Cap Gemini, Accenture* etc.).

He is an **adjunct faculty** for **Indian Institute of Management**, Bangalore and **Indian School of Business**, Hyderabad. Both are ranked among **top 30** business schools in the world.

He has co-authored the best-selling book [Machine Learning using Python](#)

He has published the following machine learning cases in (HBR) **Harvard Business Publishing**:

1. [Customer Analytics at Big Basket - Product Recommendations](#)
2. [Improving Lead Generation at Eureka Forbes Using Machine Learning Algorithms](#)

The screenshot shows the Amazon product page for the book "Machine Learning using Python Paperback". The book cover features a network diagram with orange and red nodes and edges. The title "Machine Learning using Python" is prominently displayed in white and orange text. The authors' names, "Manaranjan Pradhan | U Dinesh Kumar", are listed below the title. The Wiley logo is visible at the bottom of the cover. To the right of the cover, the product details are listed: "Machine Learning using Python Paperback – 1 January 2019" by "U Dinesh Kumar Manaranjan Pradhan (Author)". It has a 4.5-star rating from 120 ratings and is labeled as a "#1 Best Seller" in Artificial Intelligence. Two purchase options are shown: "Kindle Edition" for ₹ 434.40 and "Paperback" for ₹ 543.00 with Prime. A "FREE delivery: Wednesday, July 22" offer is also present, along with a link to "Order within 10 hrs and 5 mins Details". At the bottom, there are icons for Kindle, a book, and a truck.



## Naveen Kumar Bhansali

### **Industry Experience:**

- Co-founder - CTO & AI Head at BlitzAI (No-code AI platform)
- 18 years of industry experience as AI advisory consultant, AI solution architect and engineer, data scientist, big data architect.
- Headed AI and Big Data multi-million-dollar projects in Latin America, EMEA, USA and Southeast Asia for EMC Technologies.
- Delivered production grade solutions for clients across the globe.
  - Telecom: TIM Brazil, Hawaiian Telecom USA, MTN Nigeria
  - Finance: AMEX USA, Itau Brazil, Bank of Thailand, Bank of Ayudhya (Thailand)
  - Insurance: FWD Hong Kong
  - Retail: ASOS UK, StyleUnion India
  - Manufacturing: Embraer Brazil
  - MedTech: Exdion India

### **Academic Experience:**

- Adjunct Faculty for more than 10 years at IIM Bangalore (Ranked among top 30 business schools in the world).
- IIMBx Instructor for “AI for Managers”, covering Deep Learning and Generative AI. (<https://iimbx.iimb.ac.in/ai-for-managers/>)
- CFA Level 2
- Vice President of Analytical Society of India.
- Conducted hundreds of workshops and trained 1000s on Artificial Intelligence, Deep Learning etc. for corporates such as CISCO, JP Morgan, General Motors, EMC, Fidelity, Global Analytics India Pvt Ltd, etc. and several academic institutions across India.

### **Publications in Harvard Business Publishing:**

- Customer Analytics at Flipkart.com  
<https://hbsp.harvard.edu/product/IMB555-PDF-ENG>
- Breaking Barriers - Micro-mortgage Analytics:  
<https://hbsp.harvard.edu/product/IMB445-PDF-ENG>

