TECHNOLOGY FOR ALL

Curriculum for the course of

DATA SCIENCE

(with 1 month extensive hands on Internship)

Website: www.technologyforall.in

Contact us: +91 6304655320

TECHNOLOGY FOR ALL

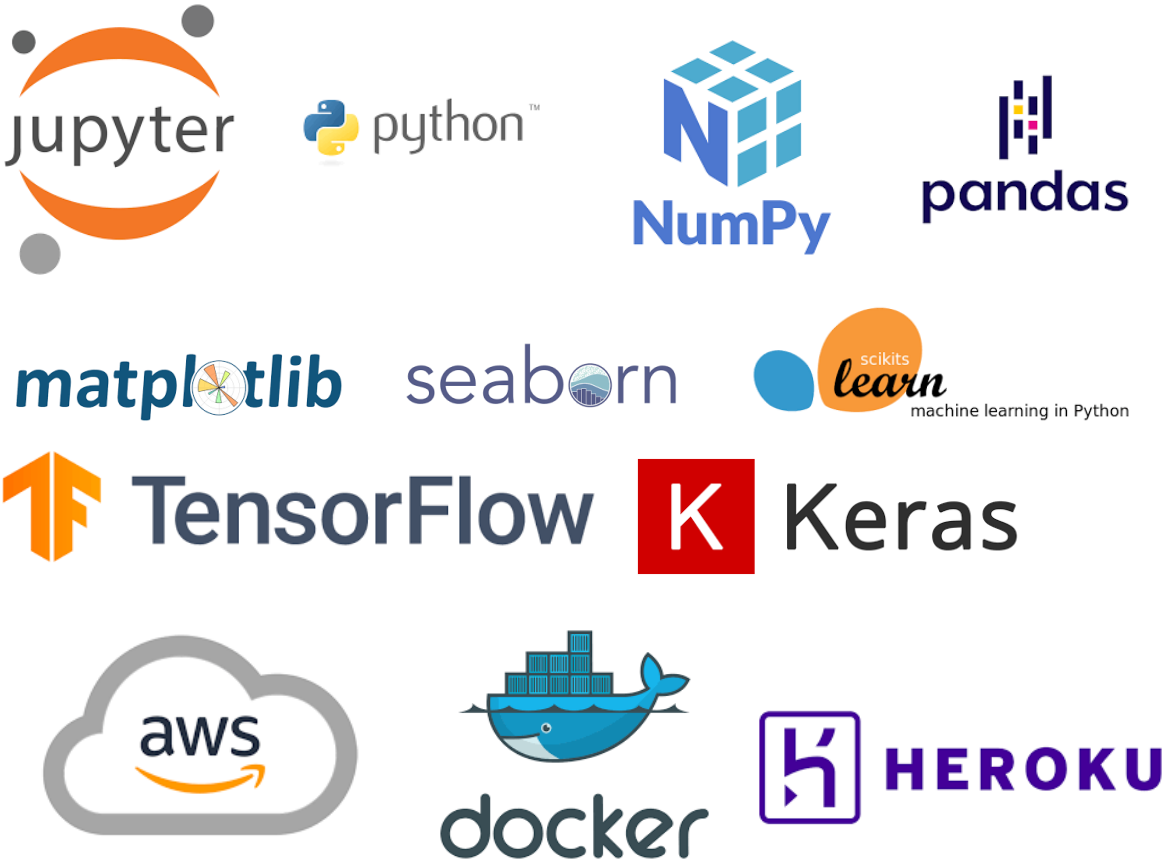
About this course

This course will serve as a comprehensive introduction to various topics in Data Science. This course is a proper blend of theory and the practical hands on session for each and every concept. Throughout the course participants will work on a complete end to end Data Science lifecycle with multiple hands-on projects.

Program Highlights

- Complete online program
- 120+ Hours of Live Lectures
- 60+ Hours of Live Handson Sessions
- 100+ Hours of Assignment and Projects
- 10+ Mini Projects
- 3+ Major Projects
- 1 Capstone Project

Programming Languages and Tools



Learning Objectives

This course will serve as a comprehensive introduction to various topics in machine learning. This course is a proper blend of theory, mathematical derivations and the practical hands on session for each and every concept. Throughout the course participants will work on numerous datasets in order to implement the concepts learned during the course.

Learning Outcomes

At the end of the course participants should be able to -

- Extract, load and transform the data.
- Scrap the data using web scraping tools and libraries
- Perform exploratory data analysis.
- Design and implement machine learning solutions to classification, regression, and clustering problems.
- Develop and Deploy Machine Learning Application
- Build end to end Deep Learning Pipelines using Tensorflow and Keras
- Work on text, image and video data
- Get handson experience on Amazon EC2 Cloud Instance

Prerequisites

There are no prerequisites required for the course. But a little knowledge of programming and high school mathematics is a benefit for candidates.

Schedule

No	Topic	Total hours
1	Python for ML – Introduction, Flow control, data structures, functions, modules, packages, regular expression, web scraping, etc	15
2	Mathematical Essentials for AI – Intro to AI, Linear Algebra, Statistics	10
3	Getting your hands dirty with data – Numpy, Pandas, Matplotlib, seaborn, Visualizations, Exploratory data analysis, Univariate, bivariate and multivariate analysis, Mini Project	20
4	Machine Learning Algorithms – Linear Regression, Logistic Regression, KNN, Performance Metrics, Decision Trees, Random Forest, GBDT, K-Means, Hierarchical Clustering, PCA, Case Study	30
5	Deep Learning – ANN, CNN, RNN, OpenCV, NLP	40
6	Machine Learning App Development(Flask Framework) and Deployment on Cloud(Heroku and AWS) and Docker	10
7	Data Analysis with SQL and Tableau	10

Module 1 - Python for Data Science

1. **Introduction**
    a. Download and install anaconda
    b. Overview of jupyter notebook
2. **Python for AI/ML**
    a. Why Python?
    b. Application areas
    c. HelloWorld in Python
    d. Python vs other languages
    e. Indentation
    f. Keywords
    g. Python Operators
3. **Data Types and Data Structures**
    a. Identifiers
    b. int, float, complex, bool
    c. Strings
    d. List, Tuple, Dictionary, Set
4. **Control Flow and Decision Statements**
    a. if, elif and else statements
    b. Nested conditional statements
    c. for loops
    d. while loops
    e. for else
    f. while else
    g. List Comprehension
5. **Functions and Modules**
    a. Introduction
    b. def keyword
    c. Scope of variables in function
    d. Defining and Calling a function
    e. Lambda, Map, Reduce and Filter
    f. Introduction to Modules and Packages
    g. import keyword
    h. Creating a Package in Python
    i. Understanding __name__ variable

TECHNOLOGY FOR ALL

Module 2 - Version Control (Git) and Flask Framework

1. **Version Control (Git)**
    a. Introduction to Version Control
    b. Introduction to Git and GitHub
    c. Why Git
    d. Centralized and Distributed Version Control System
    e. Installing Git
    f. Git Basics
    g. Forking
    h. Cloning
    i. Making changes to Local Repositories
    j. Committing
    k. Branching
    l. Collaborating with Multiple Developers
2. **Web Application Development**
    a. Introduction to Client Server
    b. Introduction to Flask Web Framework
    c. Basics of Routing
    d. Dynamic Routing
    e. Introduction to HTML
    f. HTML Forms
    g. Templates
    h. Jinja Templating
    i. Template Inheritance
    j. Flask WTForms
3. **Building Flask Application**
4. **Cloud Deployment**
    a. Introduction to Cloud
    b. Heroku Deployment
    c. AWS Deployment
5. **Introduction to Containerisation and Docker**

Module 3 - Data Analysis using Python

1. **Statistics and Probability**
    a. Fundamentals – What is Descriptive Statistics?
    b. Mean, median and mode
    c. Range, IQR, variance and standard deviation
    d. Covariance and Correlation
    e. Correlation
    f. Normal Distribution
    g. Fundamentals – Random Experiment, Event, Axioms, etc
    h. Conditional Probability
    i. Random Variable
    j. Gaussian Distribution
2. **Data Analysis with Numpy and Pandas**
    a. Intro to Numpy
    b. Creating an array
    c. Indexing and Slicing
    d. Statistical Operations using Numpy
    e. Introduction to Pandas
    f. Introduction to Series and Dataframe
    g. Working with .csv
    h. Working with .xlsx
3. **Data Manipulation with Pandas**
    a. Groupby, Pivot tables and Crosstabs
    b. Re-indexing
    c. Handling missing Values
    d. Outlier treatment
    e. Duplicates
    f. Visualization of basic plots using Pandas
4. **Data Visualization with Matplotlib and plotly**
    a. Histogram
    b. PDF
    c. Scatter Plot
    d. Pair plot
    e. Strip Plot
    f. Box Plot
    g. Violin Plot
    h. Count Plot
5. **Exploratory Data Analysis - Case Studies (Mini Projects)**
    a. Univariate analysis
    b. Bivariate Analysis
    c. Python Implementation on various datasets

Module 4 - End to End Web Scraping

1. **Regular Expressions**
    a. Understanding unstructured data
    b. Meta Characters
    c. Literals
    d. Regex in Python
    e. import re
    f. Pattern Matching
2. **Introduction to Web Scraping**
    a. requests module
    b. Installing bs4 and BeautifulSoup
    c. Loading the web pages using requests
    d. Extracting the HTML from Web Pages
    e. find() and find_all()
3. **Project on Web Scraping**
    a. Data Mining
    b. Data Preprocessing
    c. Data Visualization

Module 5 - Machine Learning

1. **Moving to Machine Learning**
    a. Why learn AI/ML?
    b. AI vs ML vs DL
    c. Applications
    d. Supervised vs Unsupervised Learning
    e. Classification vs Regression
2. **Linear Algebra**
    a. Introduction and why linear algebra?
    b. Fundamental of Vectors
    c. Fundamental of Matrices
    d. Vector Algebra
    e. Dot Product
    f. Euclidean Distance
    g. Manhattan Distance
    h. Projection
3. **Linear Regression**
    a. Introduction
    b. Equation of hyperplane
    c. Mathematical Formulation and intuition
    d. OLS Assumptions
    e. Hyperparameter Tuning
    f. Residual Analysis
    g. Polynomial Regression
    h. Non-linear transformation
    i. Feature Selection - Forward and Backward
    j. Case Study (Mini Project)
4. **K-Nearest Neighbors**
    a. Introduction
    b. Intuition
    c. Lazy Learner
    d. Deciding the number of neighbors
    e. Improving KNN performance
    f. Case Study (Mini Project)
5. **Logistic Regression**
    a. Geometric Intuition
    b. Regression vs Classification
    c. Linear Regression vs Logistic Regression
    d. Mathematical Formulation
    e. Sigmoid Function
    f. Understanding the Decision Boundary
    g. Binary vs Multiclass classification
    h. Case Study (Mini Project)
6. **Performance Measurement of Models**

a. Accuracy
b. Confusion Matrix
c. Precision and Recall
d. F1 Score
e. ROC AUC
f. Log Loss
g. R square
h. Case Study (Mini Project)

7. **Support Vector Machines**
   a. Geometric Intuition
   b. Hard and Soft Margin Classification
   c. Kernel Trick
   d. RBF-Kernel
   e. Tuning Hyperparameter
   f. Case Study (Mini Project)
8. **Decision Trees**
   a. Introduction (Rule Based Learning)
   b. How to build a decision tree
   c. Classification and Regression Trees (CART)
   d. Entropy
   e. Gini Impurity
   f. Overfitting and Pruning
   g. Tuning Hyperparameters
   h. Case Study (Mini Project)
9. **Model Tuning and Cross Validation**
   a. Introduction to Cross Validation
   b. k-fold Cross Validation
   c. Bias-Variance Trade-off
   d. Hyperparameter Tuning
10. **Ensemble Methods - Bagging and Boosting**
    a. What is an Ensemble
    b. Random Forest Algorithm
    c. Ada Boosting
    d. Gradient Boosting
    e. Tuning hyperparameter
    f. Case Study (Mini Project)
11. **Clustering - Unsupervised Learning**
    a. What is clustering?
    b. Application
    c. K – Means Algorithm
    d. Centroids
    e. Elbow Method for deciding ‘K’ in K-Means
    f. Code Sample
    g. Python Implementation
    h. Hierarchical Clustering
    i. Agglomerative Clustering

j. Interpretation of Dendrograms
k. Practical issues with Clustering Algorithms
l. Case Study (Mini Project)

12. Principal Component Analysis

a. Introduction to Dimensionality Reduction
b. Principal Components
c. EigenValues and EigenVectors
d. Transformation of Data
e. Proportion of variance explained
f. Case Study (Mini Project)

13. Glance at state of the art Deep Learning Supervised and Unsupervised Algorithms

14. Machine Learning Application Development and Deployment

15. End to End Project implementation in Machine Learning

a. Regression
b. Classification
c. Clustering
d. Recommendation Engines

Module 6 - Deep Learning (Computer Vision and NLP)

1. **Introduction to Neural Networks**
    a. Introduction to Neurons and Perceptrons
    b. Sigmoid Neuron
    c. Activation Function
    d. Cost Functions
    e. Gradient Descent and Stochastic Gradient Descent
    f. Feedforward and Backpropagation
2. **Deep Learning Frameworks**
    a. Installing Tensorflow
    b. Tensorflow and Keras
    c. Basic syntax
    d. Saving Models
    e. Tensorboard
3. **Artificial Neural Network**
    a. Intuition behind Back Propagation
    b. Computing the derivatives
    c. Training Deep Neural Networks
    d. Optimization Algorithms
    e. Activation functions and Initialization Methods
    f. Sigmoid, Tanh and ReLu
    g. Regularization Methods
    h. Overfitting and Regularisation
    i. Early Stopping
4. **Computer Vision - OpenCV**
    a. Introduction to Vision Tasks
    b. Introduction to OpenCV
    c. Working with Images
    d. Filtering
    e. Preprocessing an image with OpenCV
    f. Reshaping and Resizing
    g. Gaussian Blur, Dilation and Erosion
    h. Contours, Hull and Blobs
    i. Working with Videos
    j. Hands-on Demo
    k. Mini Project
5. **Convolution Neural Network**
    a. Introduction to CNN
    b. Convolution Operation
    c. Stride
    d. Padding
    e. Max Pooling
    f. VGG16
    g. Transfer Learning

h. AlexNet
i. GoogleNet
j. ResNet
k. Implementing CNN and Transfer Learning in Keras

6. **Industry Use Case of CNN's**
   a. Semantic Segmentation - U-Net
   b. Object Detection - YOLO and SSD
   c. RCNN - Fast and Faster
   d. Siamese Network as metric learning
   e. Hands-on demo
7. **Natural Language Processing**
   a. Tokenization
   b. Stop Words
   c. Special Characters
   d. Regular Expressions
   e. Stemming
   f. Lemmatization
   g. Bog of Word
   h. TF-IDF
   i. Word2vec
   j. Glove
   k. POS Tagger
   l. Named Entity Recognition
   m. Code Sample
   n. Sentiment analysis project
8. **Recurrent Neural Network**
   a. Introduction to RNN
   b. Architecture
   c. Types of RNN
   d. Training RNNs
   e. Bidirectional RNN
   f. LSMTs
   g. GRU
   h. Implementation in Keras
9. **Attention Models**
   a. Introduction to Encoder Decoder
   b. Auto Encoder
   c. Introduction to BERT
   d. Intuition and Application of GAN's
   e. Industry applications of GPT-1, 2 and 3
10. **Deep Learning Project Implementations**

Module 7 - Data Analysis with SQL and Tableau

1. **SQL for Data Science**
    a. Introduction to Databases
    b. Basics of SQL - Select, from and where
    c. DML, DDL and DCL
    d. Limit, offset, orderby, distinct, logical operators
    e. Joins
    f. SQL Aggregation - count, min, max, avg, sum
    g. Insert, update, delete
    h. Create, alter, add, modify, drop, truncate
    i. Grant, revoke
    j. Data manipulation and analysis using SQL
2. **Tableau for Data Science**
    a. Install Tableau for Desktop
    b. Connect Tableau to a dataset
    c. Analyze, Blend, Join and Calculate Data
    d. Tableau for Visualization
    e. Various Charts, Plots and Maps
    f. Data Hierarchies
    g. Calculated Fields
    h. Filters
    i. Creating Interactive Dashboards
    j. Adding actions to Dashboards