DataTeach.ai

DATA SCIENCE

CURRICULUM

Online/Offline

DataTeach.ai

Module 1 : PYTHON PROGRAMMING

Introduction:
- Introduction to python
- Features of python
- Advantages of python over other programming language
- Anaconda installation
- Jupyter notebook Basics

Data Types & Data Structures
- Variables
- Data Types
- Operators
- Strings
- Lists
- Sets
- Tuples
- Dictionaries

Control Flow & Conditional Statements
- Operators
- if
- elif and else statements
- range
- while loops and control flow.

- for loops and nested loops
- pass, break and continue
- Nested loops and list
- dictionary comprehensions.

Functions and Modules
- What is function and types of functions
- Code optimization and argument functions
- Lambda functions.
- Map, filter
- Manual higher order functions & nested functions
- Importing a module
- Namespace & scope of a variable using help() and dir() aliasing or renaming
- Some Important Modules In Python: math module
- random module
- datetime and os module.

Class and Objects
- Create A Class And Objects
- __init__()
- self parameter
- Class Properties
- Instance Properties & Static Properties
- Modifying Object Properties
- Delete Object
- Pass Statements
- 4 pillars of oop
- Inheritance
- Encapsulation
- Polymorphism
- Abstraction
- Multiple dispatch & abc modules.

File Handling
- Errors and Exception Handling,
- Create, Read, Write files
- Operations in File Handling.

Module 2 : DATA ANALYSIS USING PYTHON

Numpy - Numerical Python
- Arrays
- Basic Operations In Numpy
- Mathematical Functions Of Numpy
- Numpy With Images
- Advance Numpy.

Data Manipulation with Pandas
- Series and DataFrames
- Data Importing and Exporting through Excel
- CSV Files
- Data Understanding Operations
- Indexing and slicing and More filtering with Conditional Slicing
- Groupby
- Pivot table and Cross Tab
- Concatenating and Merging Joining
- Descriptive Statistics.

- Removing Duplicates
- String Manipulation
- Date Time Manipulations
- Other Forms Of Data xls
- html & json files json normalization
- Missing Data Handling mcar, mar & mnar Visualization & Imputation Of Missing Data Using Pandas
- Merges & Joins, Window Functions Statements.

Data Visualization Using Matplotlib And Seaborn
- Introduction to Matplotlib
- Plotting
- Properties of plotting
- About Subplots
- Line plots
- Pie Chart & Bar Graph.

Exploratory Data Analysis (EDA)
- What is EDA?
- Uni - Variate Analysis
- Bi - Variate Analysis
- More on Seaborn Based Plotting Including Pair Plots
- Catplot
- Heat Maps
- Count plot along with matplotlib plots.

UNSTRUCTURED DATA PROCESSING
- Regular Expressions
- Structured Data and Unstructured Data
- Literals and Meta Characters, How to Regular Expressions using Pandas?
- Inbuilt Methods
- Pattern Matching, flags.

Project (web - scraping)

Module 3 : STATISTICS

Syllabus

- Vectors
- Matrices
- Linear Equation
- Vector Norms
- Calculus
- Central Tendency
- Sample V/S Population

- Probability
- Types of Distribution
- Central Limit Theorm
- Hypothesis Testing
- Types of Test
- Anova Test
- Skewness & Kurtosis

Module 4: SQL

Syllabus

Introduction to Basic Database Concepts:
- What is Data, Field, Record and database? ,Limitations of File Management System, Basic Concepts of Advantages of DBMS, Level of abstraction, Database models, Exploring Relational DBMS, Discuss the basic design, theoretical, and physical aspects of a relational database, Understanding Client and Server, What is MySQL?

Introduction to SQL
- MySQL datatypes, Basics of Types of SQL Statements, Create and use Database, Categorize the different types of SQL statements, DDL, DML, DQL, DCL and TCL, Data types in SQL, Exploring DDL Statements on Table

Writing Basic SQL Statement
. List the capabilities of SQL SELECT statements, Generate a report of data from the output of a basic SELECT statement, Select All Columns, Select Specific Columns, Use Column Heading Defaults, Use Arithmetic Operators, Understand Operator Precedence, Learn the DESCRIBE command to display the table structure, Using Parentheses,

- Defining a Null, Defining a Column Alias, Concatenation Operator, Literal Character Strings, Eliminating Duplicate Rows

Restricting and Sorting Data
- Limiting Rows Using a Selection, Limiting the Rows Selected, Using the WHERE Clause, Character Strings and Dates, Using Comparison Conditions, BETWEEN Condition, IN Condition, LIKE Condition, NULL Conditions, Logical Conditions, AND Operator, Using the OR Operator, Using the NOT Operator, Rules of Precedence, ORDER BY Clause, Sorting in Descending Order, Sorting by Column Alias, Sorting by Multiple Columns

Advance DDL Commands
- Normalization, Create Database objects, Alter Table Statements, Drop Table Statements, Various Constraints, Commit, Rollback, Savepoint, Creating Views.

Working on DML statements
Data Manipulation Language, Adding a New Row to a Table, The INSERT Statement Syntax, Inserting New Rows, Inserting Rows with Null Values, Inserting Special Values, Inserting Specific Date Values.

- Creating a Script, Copying Rows from another Table

Use of built-in function in SQL
- Conversion Function, Logical Functions, Math Function, Aggregate Functions, String Functions, Date Functions

Working on multiple tables and Retrieve records from multiple tables
- Self-Join, Inner Join, Left Join, Right Join, Cross Join

Working on subqueries
- Using a Subquery to Solve a Problem, Subquery Syntax, Using a Subquery, Guidelines for Using Subqueries, Types of Subqueries, Single-Row Subqueries, Executing Single-Row Subqueries, Using Group Functions in a Subquery, The HAVING Clause with Subqueries, Multiple-Row Subqueries, Using the ANY Operator in Multiple-Row Subqueries, Using the ALL Operator in Multiple-Row Subqueries, Null Values in a Subquery.

Python with MYSQL
- Regex with MYSQL, Regex with Python, Python MySQL Limit: Limit the Result, Start From Another Position, Quick intro of SQLIA, Import, Export, Overview of Cursor and Trigger

DataTeach.ai

Module 5 : POWER BI

Syllabus

- Intro to Power BI
- Basic Charts in Power BI
- Maps in Power BI
- Tabular Data in Power BI
- Cards in Power BI
- Filters in Power BI
- Slicers in Power BI
- Advance Chart in Power BI

- Inserting Object in Power BI
- BookMark Action in Power BI
- Creating Report in Power BI
- Publishing Report in Power BI
- Service Dashboard in Power BI
- Refreshing Data in Power BI and Scheduling

- Using Text Functions, Date Functions, Number Functions in Power Query
- Appending and Merging Sheet
- Columns from example in power query, Conditional ColumnsN
- Data Modelling in Power BI
- Connecting to MySql

Module 6 : MACHINE LEARNING

Syllabus

Introduction
- Data Cleaning,
- Data Pre-processing
- Data Modelling
- Overfitting
- Underfitting
- Bias
- Variance
- Trade-off

Supervised Learning:
Regression Methods:
- Simple Linear Regression,
- Multiple Linear Regression
- Polynomial Regression
- Ridge Regression
- Lasso Regression.

Regression Metrics:
- Mean-Square-Error
- Root-Mean-Square-Error
- Mean-Absolute-Error
- R2-Socre.

classification Methods:
- Logistic Regression with Single Variable
- Logistic Regression with Multiple Variable
- KNN (K – Nearest – Neighbors)
- Naïve Bayes Classifier
- Support Vector Machines
- Decision Tree Classifier
- Random Forest Classifier.

Classification Metrics:
- Confusion Matrix (TP, FP, TN, FN)
- Accuracy-score
- Precision
- Recall, F1-score
- ROC-Curves (Receiver-Operating-Characteristics)
- Dimensionality Reduction
- Principal Component Analysis
- Anomaly detection algorithm
- Unsupervised Learning
- K-Means Clustering
- Hierarchical Clustering
- Density-based spatial clustering of applications with noise.

Module 7 : DEEP LEARNING

Syllabus

Introduction to Deep Learning
- Intro To AI
- ML AND DL
- Difference between ML and DL
- When to use ML and DL
- History Of Deep Learning
- Intro to Biological Neuron.

Neural Network Architecture and Activation Functions
- Introducing Google Colab
- Tensorflow basic syntax
- Tensorflow Graphs
- Tensorboard

Forward and Backward Propagation
- MLP Architecture
- Defining the Notation for MLP
- Working of MLP (Forward Propagation).

- How To Train Single Neuron Model
- Backpropagation -1 (chain rule)
- Backpropagation -2 (chain rule+ memorization
- Hyperparameter In MLP
- Bias and Variance Trade-off In MLP
- Why Deep Neural Network Failed
- Vanishing Gradient Problem
- Exploding Gradient Problem
- Activation Function -3 (ReLU and ReLU Variants Linear and Non Linear Variants) [Leaky ReLU, parametric ReLU, ELU, SELU]
- Dropouts
- Weight Initialization Techniques (pros and cons)
- Batch Normalization
- Early Stopping, Tensor Board.

Optimizers
- Convex Function And Non Convex Functions
- Saddle Point
- Random Forest Classifier
- SGD
- NAG
- Rmsprop Ada Delta
- Ada Grad
- ADAM
- NADAM.

Keras Hands-on - Regression and Classification
- Intro To Tensorflow and Keras
- Project on Classification by using MLP
- Project on Regression by using MLP.

Module 8 : COMPUTER VISION

Syllabus

Intro to Images and Image Preprocessing with OpenCV
- Intro To Images
- How Images are formed and stored in machines
- Color Spaces
- Intro To OpenCv
- read, write, save image
- Converting to Different Color Spaces
- Building Histograms for Images.

Image Preprocessing with OpenCV
- Read videos
- Capturing images with web camera
- Manipulating videos with opencv
- Drawing on images and videos
- Bitwise Operators On Images and Videos
- Affine and Non-Affine Transformation.

Intro to Convolutional Neural Network

Image Classification Case Study
- Face Mask Detection
- Bone fracture Multi region detection.

CNN Architecture
- Padding
- Stride
- Pooling
- LeNet5 Alex
- Net
- Vgg 16 and Vgg 19
- Inception Net ResNet
- Xception
- Mobile Net, Efficient Net
- Pre trained Model Introduction.

Transfer Learning
- Intro To Transfer Learning
- Transfer learning Concepts (When and Why)
- Transfer Learning Coding
- Hyper Parameter Tuning [Random Search

Hyperband, Bayesian optimization.

Case Study with Transfer Learning
- Plant Diseases Prediction using Transfer Learning
- Cifar using Transfer Learning, Improving Face Mask Detection Model using Transfer Learning.

Object Detection
- Intro To object Detection
- R-CNN, Fast R-CNN
- Faster R-CNN # Show why Faster R CNN is faster than R CNN (no Need of Maths)
- Intro to Yolo
- Yolo Algorithm (How it works) - More Detail on YOLO
- Implementation of Yolo V7 / V8 using Ultralytics.

YOLO - Case Study
- Helmet Detection using Yolo.

Module 9 : NATURAL LANGUAGE PROCESSING

Syllabus

Introduction to text and Text Preprocessing with nltk and spacy
- Intro to NLP
- Text Preprocessing Steps
- Tokenization
- Special Character
- Stop words, Stemming & Lemmatization.

Vectorization Techniques
- BOW
- TF-IDFCoding for BOW and TF-IDF using nltk, Word2Vec
- Glove
- How Word2Vec algorithm works (Skip-Gram & CBOW)
- FastText.

Project - Text Classification
- Word2Vec, Glove & FastText.

RNNs
- Intro to RNN
- Why RNN ?
- How RNN Works
- Trainning RNN
- Types of RNN.

Project - Sequence Tagging
- NER and POS Tagging case study.

LSTMs
- Intro to LSTM
- Why LSTM
- LSTM algorthim
- Grus, Bi-Directional RN, Understanding of working of Image captioning.

Auto Encoders
- Encoder Decoder Architecture
- Introduction to autoencoders
- Types of autoencoders.

Project - Auto Encoders
- Case study for Encoder decoder and autoencoder for image compression and reconstruction on MNIST Images.

Transformer and Attention
- Intro to Transformers and Attention Models.
- How does Transformers works, How does Attention works
- Coding For Transformers and Attention Models.

BERT
- Intro to BERT
- How does BERT works
- Coding For Transformers and Attention Models.

DataTeach.ai

Module 10 : ## GEN AI

Intro To Gen AI
- Introduction to Generative AI
- Overview of generative AI technologies
- Applications and case studies across industries.

Intro To LLM
- History of NLP
- Into to large language Models.
- What is Large Language Model.

- Types of Large Language Model.

Prompt Engineering and Working with LLM
- Intro To Open AI
- Utilizing OpenAI APIs
- Setting up and authenticating API usage
- Practical exercises using GPT-3/GPT-4 for text generation
- Understanding DALL-E.

Open AI
- Intro To Open AI
- Utilizing OpenAI APIs
- Setting up and authenticating API usage
- Practical exercises using GPT-3/GPT-4 for text generation.

TOOLS COVERED

python	statistics	scikit learn	TensorFlow	NumPy pandas	plotly	NLTK	Keras
matplotlib	Power BI	RegEx	OpenCV	seaborn	SQL	jupyter	OpenAI
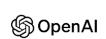
KEY HIGHLIGHTS OF THE TRAINING

100+ hours of learning	20+ Use Cases
Instructor-Led LIVE Sessions	Mock Interviews
Life-time LMS Access	Learners Community Access
LIVE Doubt Resolution	Guaranteed Internship
Topic-wise Assignments	Soft Skills
Weekly Quizzes	Resume Building
Monthly Assessments	100% Job Assistance

LMS Portal Login:
www.datateach.ai

+91 988 594 6789   www.datateach.ai   info@datateach.ai