CETPA®

TRAINING | DEVELOPMENT | PLACEMENT

www.cetpainfotech.com

Data Science Using Python

Our Partners

Microsoft Partner Silver Learning

AUTODESK.

Panasonic

nuvoTon Training Partner

Wonderware by Schneider Electric

CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

CURRICULUM OF CORE & ADVANCED PYTHON

GETTING STARTED

- History & need of Python
- Application of Python
- Advantages of Python
- Disadvantages of Python
- Installing Python
- Program structure
- Interactive Shell
- Executable or script files.
- User Interface or IDE

PYTHON FUNDAMENTALS

- Working with Interactive mode
- Working with Script mode
- Python Character Set
- Python Tokens, Keywords, Identifiers, Literals, Operators
- Variables and Assignments
- Input and Output in Python

DATA HANDLING

- Data Types
- Numbers
- Strings
- Lists
- Tuples
- Dictionary
- Set
- Frozenset
- Bool
- Mutable and Immutable

STRING MANIPULATION

- Introduction to Python String
- Accessing Individual Elements
- String Operators
- String Slices
- String Functions and Methods

LIST MANIPULATION

- Introduction to Python List

CURRICULUM OF Data Science Using Python

- Creating List
- Accessing List
- Joining List
- Replicating List
- List Slicing

TUPLES

- Introduction to Tuple
- Creating Tuples
- Accessing Tuples
- Joining Tuples
- Replicating Tuples
- Tuple Slicing

DICTIONARIES

- Introduction to Dictionary
- Accessing values in dictionaries
- Working with dictionaries
- Properties

SET AND FROZENSET

- Introduction to Set and Frozenset
- Creating Set and Frozenset
- Accessing and Joining
- Replicating and Slicing

OPERATORS

- Arithmetic Operators
- Relational Operators
- Logical Operators
- Membership Operators
- Identity Operators
- Bitwise Operators
- Assignment Operators
- Operators Precedence
- Evaluating Expression
- Type Casting

PROGRAM CONTROL FLOW

- Conditional Statements
- The if Statement

CURRICULUM OF Data Science Using Python

The if-else Statement
The if-elif Statement
Nested if Statements
Python Indentation
Looping and Iteration
The For Loop
The While Loop
Loop else Statement
Nested Loops
Break and Continue
The Range Function
Introduction to range()
Types of range() function Use of range() function

INTRODUCTION TO FUNCTIONS

Built-In Functions

Introduction to Functions
Using a Functions
Python Function Types
Structure of Python Functions
E.g. - map, zip, reduce, filter, any, chr, ord, sorted, globals, locals, all, etc.

User Defined Functions

Structure of a Python Program w.r.t. UDF
Types of Functions
Invoking UDF
Flow of Execution
Arguments and Parameters
Default Arguments, Named Arguments
Scope of Variables
Lambda function

Recursion Function

Use of recursion function

MODULES AND PACKAGES

Built-in Modules

Importing Modules in Python Programs
Working with Random Modules
E.g. - builtins, os, time, datetime, calendar, twilio, smtp, pillow.

CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

User Defined Modules

- Structure of Python Modules

FILE OPERATIONS

Text and Bytes files

- Opening a file
- Reading and Writing Files
- Other File tools

JSON/PICKLE

FORMAT CLASSES AND OBJECTS

- Classes as User Defined Data Type
- Objects as Instances of Classes
- Creating Class and Objects
- Creating Objects By Passing
- Values Variables & Methods

EXCEPTION HANDLING

- Default Exception and Errors
- Catching Exceptions
- Raise an exception
- Try ...except statement
- Raise, Assert, Finally blocks
- User defined exception

INTRODUCTION TO OOPS

- Procedural Vs Modular Programming
- The Object Oriented Programming
- Data Abstraction
- Data Hiding
- Encapsulation
- Inheritance
- Polymorphism
- Generators
- Iterators

CETPA®

TRAINING | DEVELOPMENT | PLACEMENT

www.cetpainfotech.com

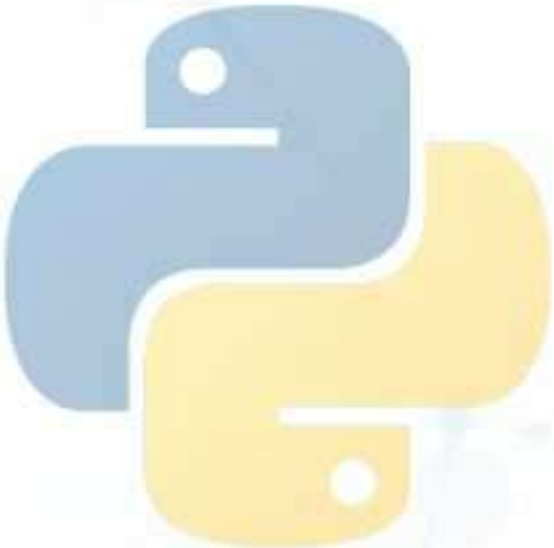
CURRICULUM OF Data Science Using Python

DATABASE

- SIntroduction to MySQL
- PYMYSQL Connections
- Executing queries
- Transaction Handling error

GUI PROGRAMMING

- Introduction
- Tkinter programming
- Tkinter widgets
- Frame
- Button
- Label
- Entry
- All Widget



CORPORATE OFFICE: D-58, Sector-2, Near Red FM. Noida -201301, Uttar Pradesh
Contact Us: +91-9212172602 , 0120-4535353

HEAD OFFICE: 200 Purwavali, 2nd Floor, (Opp. Railway Ticket Agency), Railway Road , Ganeshpur, Roorkee – 247667, Ph.No.: +91-9219602769, +91-9219602743

BRANCH OFFICE: 401 A, 4th Floor, Lekhraj Khazana, Faizabad Road, Indira Nagar, Lucknow-220616 (U.P.) Ph. No: +91-522-6590802, +91-9258017974, +91-8800340777

BRANCH OFFICE: 105, Mohit Vihar, Near Kamla Palace, GMS Road, Dehradun-248001, UK
Contact: +91-9219602771, 0135-6006070

Our Partners

Microsoft Partner | AUTODESK | Panasonic | CompTIA Get IT Certified

CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

REVISITING PYTHON
- Revisiting Python
- List and dictionary comprehension
- Programming assignment

INTRODUCTION TO DATA ANALYTICS
- Why Analytics?
- Traditional Data Management
- Types of Analytics
- Dimensions and measures
- Why learn Python for data analysis?

LIBRARIES FOR DATA ANALYTICS
- Anaconda
- Numpy, Scipy, Pandas
- Matplotlib, Seaborn

STATISTICS:
- Mean, Median, Mode
- Z-scores
- Bias -variance dichotomy
- Sampling and t-tests
- Sample vs Population statistics
- Random Variables
- Probability distribution function
- Expected value
- Binomial Distributions
- Normal Distributions
- Central limit Theorem
- Hypothesis testing
- Z-Stats vs T-stats
- Type 1 type 2 error
- Chi Square test
- ANOVA test and F-stats

JUPYTER NOTEBOOK
- Create Documentation
- Code mode
- Markdown mode
- Heading mode

CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

NUMPY:

Creating NumPy arrays
Indexing and slicing in NumPy
Downloading and parsing data
Creating multidimensional arrays
NumPy Data types
Array tributes
Indexing and Slicing
Creating array views copies
Manipulating array shapes I/O

SCIPY:

Introduction to SciPy
Create function
modules of SciPy

PANDAS:

Using multilevel series
Series and Data Frames
Grouping, aggregating
Merge DataFrames
Generate summary tables
Group data into logical pieces
Manipulate dates
Creating metrics for analysis
Data wrangling
Merging and joining
Analytics Vidhya dataset- Loan Prediction Problem
Data Mugging using Pandas
Building a Predictive Model

MATPLOTLIB:

Scatter plot
Bar charts, histogram
Stack charts
Legend title Style
Figures and subplots
Plotting function in pandas
Labelling and arranging figures
Save plots

CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

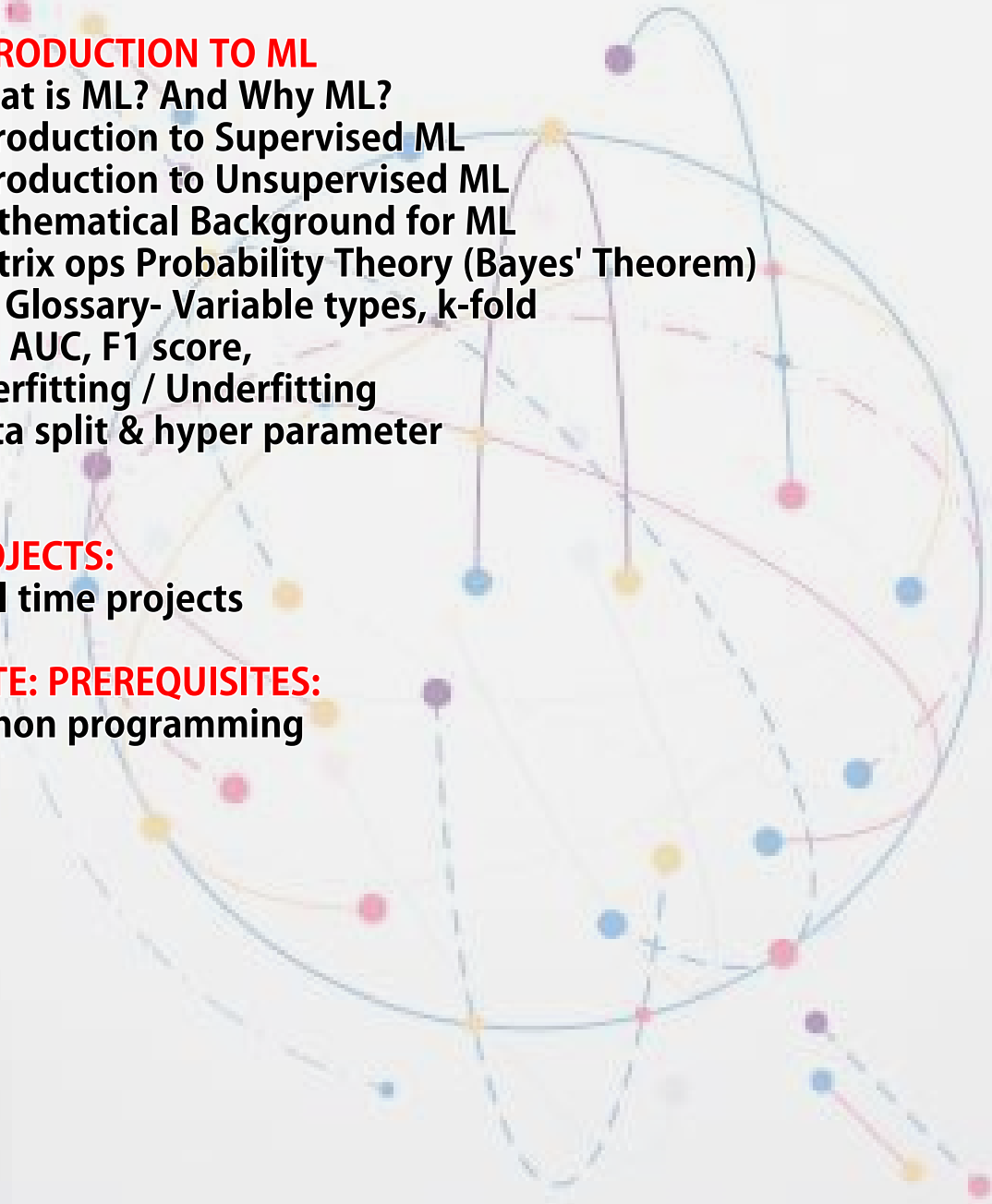
SEABORN:
Style functions, Color palettes
Distribution and Categorical plots
Regression plots
Axis grid objects

WEB SCRAPING:
Scraping Webpages
Beautifulsoup package
Real time project

INTRODUCTION TO ML
What is ML? And Why ML?
Introduction to Supervised ML
Introduction to Unsupervised ML
Mathematical Background for ML
Matrix ops Probability Theory (Bayes' Theorem)
ML Glossary- Variable types, k-fold
CV, AUC, F1 score,
Overfitting / Underfitting
Data split & hyper parameter

PROJECTS:
Real time projects

NOTE: PREREQUISITES:
Python programming



CETPA®
TRAINING | DEVELOPMENT | PLACEMENT
www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

Machine Learning

Introduction to ML
What is ML? Why ML?
Introduction to Supervised ML
Introduction to Unsupervised ML
Difference Between AI|DL|ML
Application and Use.

Tools required for development - Anaconda, Jupyter NB/Google Colab/Spyder

ML libraries
Numpy: Introduction to Numpy
pandas:Introduction| DataFrame| Loading datasets| Loading data from database| pandas Operation.
Matplotlib: Introduction|Line Chart|Pie Chart| Scatter Plot|Bar chart |Histogram.
Sklearn: : Introduction|Sklearn-API| Statsmodels.api.
ML Glossary-
Variable types, k-fold CV, AUC , F1 score,Overfitting/Underfitting, Generalization,ROC|Confusion matrix
Mathematical Background for ML- Matrix ops
Probability Theory (Bayes' Theorem)
Statistical knowledge for ML- Mean, Median, Mode, Z-scores, bias -variance dichotomy
Exploratory Data analysis using Visualisation
Scikit-learn Library for ML
Code Exercises

Steps of Machine Learning

Data Collection. The quantity & quality of your data dictate how accurate our model is. ...
Data Preparation. Wrangle data and prepare it for training|Data wrangling using Pandas|Preprocessing

data and feature engineering| Data split

Choose a Model.
Train the Model. ...

CURRICULUM OF Data Science Using Python

Evaluate the Model. ...
6 - Parameter Tuning| hyper parameter training
Make Predictions.

Supervised Learning

Introduction|Maths behind Supervised Machine Learning and Algo.

Regression:

Linear Regression
Multi-Linear Regression
Lasso/Rigde
Decision Tree Regressor
Support Vector Regressor

Classification

Logistic Regression
KNN- K Nearest Neighbors
Support Vector Classifier(SVM-SVC)
Decision Tree Classifier(DTC)
Random Forest
Naïve Bayes
Ensemble Learning

Unsupervised Machine Learning

Clustering :

Introduction:
Mathematics behind Clustering

k-means clustering

Implementation of K-mean Clustering

H-clustering

Implementation of H-clustering
Code Exercises

Assocation Rule:

Apiori rule

Dimensionality Reduction

Principle Component Analysis(PCA) .

CETPA®

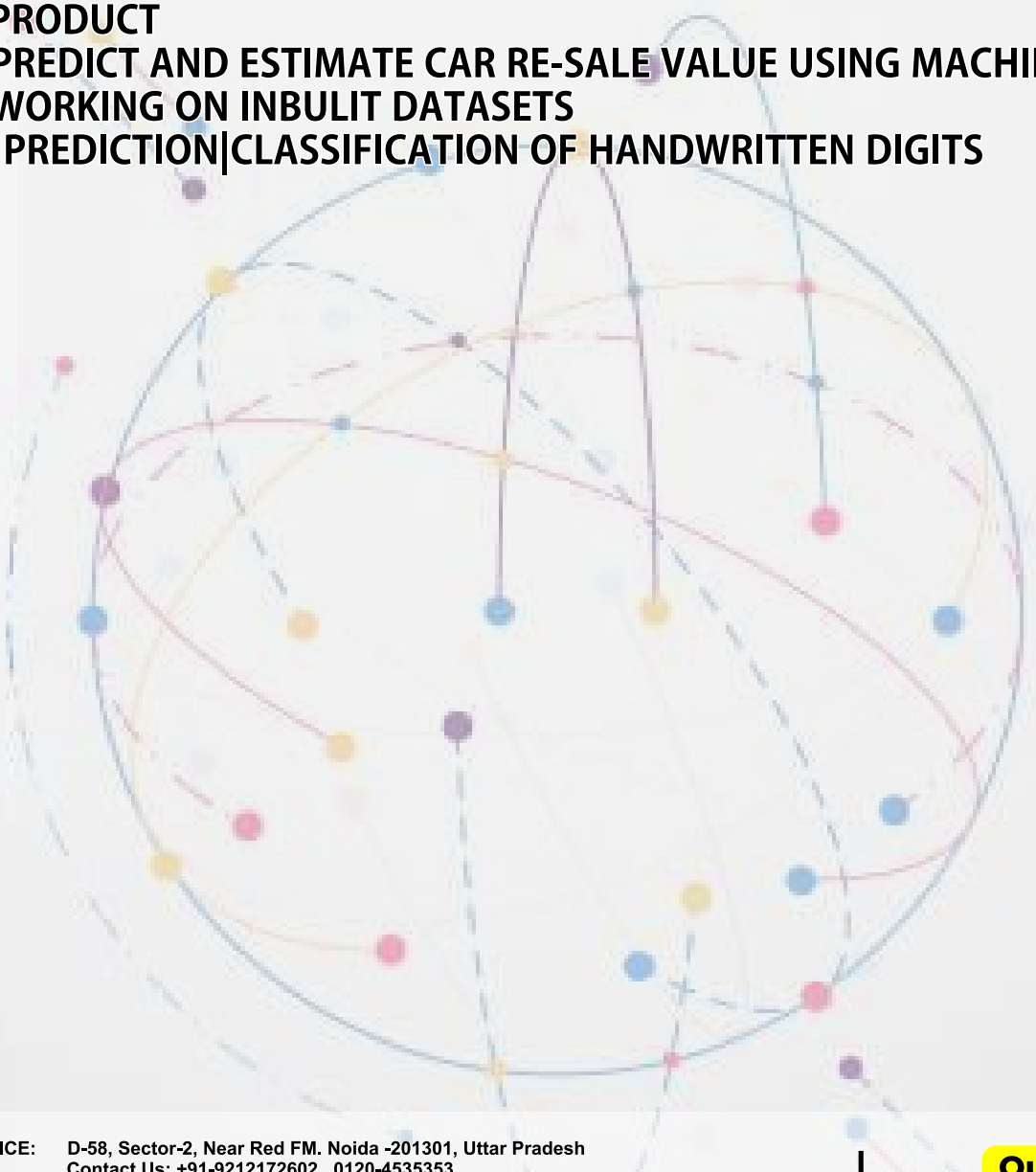
TRAINING | DEVELOPMENT | PLACEMENT

www.cetpainfotech.com

CURRICULUM OF Data Science Using Python

CAPSTONE PROJECT

1) IBM ATTRITION RATE PREDICTION USING MACHINE LEARNING
2) COVD-19 PATIENT OUTCOME PREDICTION USING ML
3) ESTIMATE THE ONLINE SALES OF A E-COMMERCE RETAIL FIRM USING ML
4) GUI BASED MACHINE LEARNING APPLICATION TO CLASSIFY THE PLANT SPECIES OF IRIS FLOWER
5) PREDICT THE CHURN RATE IN A TELECOM COMPANY USING ML
6) MALL CUSTOMER SEGMENTATION USING ML
7) MARKET BASKET ANALYSIS AND ASSIT A SHOPPING MALL TO STACK PRODUCT
8) PREDICT AND ESTIMATE CAR RE-SALE VALUE USING MACHINE LEARNING
9) WORKING ON INBULIT DATASETS
10) PREDICTION|CLASSIFICATION OF HANDWRITTEN DIGITS



CORPORATE OFFICE: D-58, Sector-2, Near Red FM. Noida -201301, Uttar Pradesh
Contact Us: +91-9212172602 , 0120-4535353

HEAD OFFICE: 200 Purwavali, 2nd Floor, (Opp. Railway Ticket Agency), Railway Road , Ganeshpur, Roorkee – 247667, Ph.No.: +91-9219602769, +91-9219602743

BRANCH OFFICE: 401 A, 4th Floor, Lekhraj Khazana, Faizabad Road, Indira Nagar, Lucknow-220616 (U.P.) Ph. No: +91-522-6590802, +91-9258017974, +91-8800340777

BRANCH OFFICE: 105, Mohit Vihar, Near Kamla Palace, GMS Road, Dehradun-248001, UK
Contact: +91-9219602771, 0135-6006070

Our Partners

Microsoft Partner | AUTODESK | Panasonic | CompTIA Get IT Certified