C-TAG

Coding Training Academy By Gandhis

DATA SCIENCE WITH PYTHON

SYLLABUS

	DEEP DIVE PYTHON
1	**Introduction to Python** Overview of Python The Companies using Python Different Applications where Python is Used Discuss Python Scripts on UNIX/Windows Values, Types, Variables Operands and Expressions Conditional Statements Loops Command Line Arguments Writing to the Screen
2	**Sequences and File Operations** Strings and related operations Tuples and related operations Lists and related operations Dictionaries and related operations Sets and related operations
3	**Deep Dive – Functions, OOPs, Modules, Errors and Exceptions** Functions Function Parameters Global Variables Variable Scope and Returning Values Lambda Functions Object Oriented Concepts Standard Libraries Modules Used in Python The Import Statements Module Search Path Package Installation Ways Errors and Exception Handling Handling Multiple Exceptions
	DATA ANALYSIS
4	**Introduction to NumPy, Pandas and Matplotlib** Data Analysis NumPy - arrays Operations on arrays Indexing, slicing, and iterating Reading and writing arrays on files Pandas - data structures & index operations Reading and Writing data from Excel/CSV formats into Pandas

	Metadata for imported Datasets Matplotlib library Grids, axes, plots Markers, colors, fonts, and styling Types of plots - bar graphs, pie charts, histograms Contour plots
5	**Data Manipulation** Basic Functionalities of a data object Merging of Data objects Concatenation of data objects Types of Joins on data objects Exploring a Dataset Analyzing a dataset
	MACHINE LEARNING
6	**Introduction to Machine Learning with Python** Python Revision (numpy, Pandas, scikit learn, matplotlib) What is Machine Learning? Machine Learning Use-Cases Machine Learning Process Flow Machine Learning Categories Linear regression
7	**Supervised Learning - I** What are Classification and its use cases? What is a Decision Tree? Algorithm for Decision Tree Induction Creating a Perfect Decision Tree Confusion Matrix What is Random Forest?
8	**Dimensionality Reduction** Introduction to Dimensionality Why Dimensionality Reduction PCA Factor Analysis Scaling dimensional model LDA
9	**Supervised Learning - II** What is Naïve Bayes? How Naïve Bayes works? Implementing Naïve Bayes Classifier What is a Support Vector Machine? Illustrate how Support Vector Machine works. Hyperparameter Optimization Grid Search vs. Random Search Implementation of Support Vector Machine for Classification

	Implementation of Naïve Bayes, SVM
10	**Unsupervised Learning** What is Clustering & its Use Cases? What is K-means Clustering? How does the K-means algorithm works? How to do optimal clustering What is C-means Clustering? What is Hierarchical Clustering? How does Hierarchical Clustering work?
11	**Association Rules Mining and Recommendation Systems** What are Association Rules? Association Rule Parameters Calculating Association Rule Parameters Recommendation Engines How do Recommendation Engines work? Collaborative Filtering Content-Based Filtering
12	**Reinforcement Learning** What is Reinforcement Learning Why Reinforcement Learning Elements of Reinforcement Learning Exploration vs. Exploitation dilemma Epsilon Greedy Algorithm Markov Decision Process (MDP) Q values and V values Q – Learning Values Calculating Reward Discounted Reward Calculating Optimal quantities Implementing Q Learning Setting up an Optimal Action
	TIME SERIES
13	**Time Series Analysis** What is Time Series Analysis? Importance of TSA Components of TSA White Noise AR model MA model ARMA model ARIMA model Stationarity ACF & PACF

14	**Model Selection and Boosting** What is Model Selection? Need for Model Selection Cross – Validation What is Boosting? How do Boosting Algorithms work? Types of Boosting Algorithms Adaptive Boosting
15	**Statistical Foundations (Self-Paced)** What is Exploratory Data Analysis? EDA Techniques EDA Classification Univariate Non-graphical EDA Univariate Graphical EDA Multivariate Non-graphical EDA Multivariate Graphical EDA Heat Maps
	PLOTTING
16	**Advanced Visualizations (Self-Paced)** Trend lines Reference lines Forecasting Clustering Geographic Maps Using charts effectively Dashboards Story Points Visual best practices Publish to Tableau Online
17	**Data Connection and Visualization in Tableau (Self-Paced)** Data Visualization Business Intelligence tools VizQL Technology Connect to data from the File Connect to data from the Database Basic Charts Chart Operations Combining Data Calculations
	DATA SCIENCE
18	**Introduction to Data Science** What is Data Science? What does Data Science involve? Era of Data Science Business Intelligence vs Data Science

	Life cycle of Data Science Tools of Data Science Introduction to Big Data and Hadoop Introduction to R Introduction to Spark Introduction to Machine Learning
19	**Statistical Inference** What is Statistical Inference? Terminologies of Statistics Measures of Centers Measures of Spread Probability Normal Distribution Binary Distribution
20	**Data Extraction, Wrangling and Exploration** Data Analysis Pipeline What is Data Extraction Types of Data Raw and Processed Data Data Wrangling Exploratory Data Analysis Visualization of Data
21	**Introduction to Machine Learning** What is Machine Learning? Machine Learning Use-Cases Machine Learning Process Flow Machine Learning Categories Supervised Learning algorithm: Linear Regression and Logistic Regression
22	**Classification Techniques** What are classification and its use cases? What is Decision Tree? Algorithm for Decision Tree Induction Creating a Perfect Decision Tree Confusion Matrix What is Random Forest? What is Navies Bayes? Support Vector Machine: Classification
23	**Unsupervised Learning** What is Clustering & its use cases What is K-means Clustering? What is C-means Clustering? What is Canopy Clustering What is Hierarchical Clustering?
24	**Recommender Engines**

	What is Association Rules & its Use Cases? What is Recommendation Engine & its Workings? Types of Recommendations User-Based Recommendation Item-Based Recommendation Difference: User-Based and Item-Based Recommendation Recommendation Use Cases
25	**Text Mining** The concepts of text-mining Use cases Text Mining Algorithms Quantifying text TF-IDF Beyond TF-IDF
26	**Time Series** What is Time Series data? Time Series variables Different components of Time Series data Visualize the data to identify Time Series Components Implement ARIMA model for forecasting Exponential smoothing models Identifying different time series scenario based on which different Exponential Smoothing model can be applied Implement respective ETS model for forecasting
27	**Deep Learning** Reinforced Learning Reinforcement learning Process Flow Reinforced Learning Use cases Deep Learning Biological Neural Networks Understand Artificial Neural Networks Building an Artificial Neural Network How ANN works Important Terminologies of ANN's