Data Science in Python
Advanced Topics

In this age of big data, companies across the globe use Python to sift through the avalanche of information at their disposal. By storing, filtering, managing, and manipulating data in Python, you can give your company a competitive edge & boost your career to the next level. This course will give you a robust grounding in all aspects of data science, from statistical modelling to visualization to machine learning. This course deep dive into the statistical modelling features of Python and gives you grounding in Python Data Science, from carrying out simple visualizations and data explorations to statistical analysis to machine learning to finally implementing simple deep learning-based models using Python. At the end of this course you will understand deep concepts like statistical modelling in Python's Statsmodels package and the difference between statistics and machine learning

You Must Know!

Duration:
40 Hours

Who should attend?
Anyone Who Wishes To Learn Practical Data Science Using Python, Learn How To Implement Machine Learning Algorithms Using Python, or Looking To Get Started In Deep Learning Using Python. Prior Knowledge Of Python Will Be Useful But NOT Necessary

Main Topics:
- Data Visualizations
- Statistical Data Analysis
- Machine Learning for Data Science
- Supervised Learning
- ANN & DL

Course modules

Module 1 – Introduction to Data Visualizations

- o What is Data Visualization?
- o Some Theoretical Principles Behind Data Visualization
- o Histograms-Visualize the Distribution of Continuous Numerical Variables
- o Boxplots-Visualize the Distribution of Continuous Numerical Variables
- o Scatter Plot-Visualize the Relationship Between 2 Continuous Variables
- o Barplot
- o Pie Chart
- o Line Chart

Module 2 – Statistical Data Analysis

- o What is Statistical Data Analysis?
- o Pointers on Collecting Data for Statistical Studies
- o Pointers on Exploring Quantitative Data
- o Explore the Quantitative Data: Descriptive Statistics
- o Grouping & Summarizing Data by Categories
- o Visualize Descriptive Statistics-Boxplots
- o Common Terms Relating to Descriptive Statistics
- o Data Distribution-Normal Distribution
- o Check for Normal Distribution
- o Standard Normal Distribution and Z-scores
- o Confidence Interval-Theory
- o Confidence Interval-Calculation

Module 3 – Statistical Inference & Relationship Between Variables

- o What is Hypothesis Testing?
- o Test the Difference Between Two Groups
- o Test the Difference Between More Than Two Groups
- o Explore the Relationship Between Two Quantitative Variables
- o Correlation Analysis
- o Linear Regression-Theory
- o Linear Regression-Implementation in Python
- o Conditions of Linear Regression
- o Conditions of Linear Regression-Check in Python

- o Polynomial Regression
- o GLM: Generalized Linear Model
- o Logistic Regression

Module 4 – Machine Learning for Data Science

- o How is Machine Learning Different from Statistical Data Analysis?
- o What is Machine Learning (ML) About? Some Theoretical Pointers
- o Unsupervised Classification – Some Basic Ideas
- o KMeans-theory
- o KMeans-implementation on the iris data
- o Quantifying KMeans Clustering Performance
- o KMeans Clustering with Real Data
- o How Do We Select the Number of Clusters?
- o Hierarchical Clustering-theory
- o Hierarchical Clustering-practical
- o Principal Component Analysis (PCA)-Theory
- o Principal Component Analysis (PCA)-Practical Implementation

Module 5 – Supervised Learning

- o What is This Section About?
- o Data Preparation for Supervised Learning
- o Pointers on Evaluating the Accuracy of Classification and Regression Modelling
- o Using Logistic Regression as a Classification Model
- o RF-Classification
- o RF-Regression
- o SVM-Linear Classification
- o SVM-Non Linear Classification
- o Support Vector Regression
- o knn-Classification
- o knn-Regression
- o Gradient Boosting-classification
- o Gradient Boosting-regression
- o Voting Classifier

Module 6 – Artificial Neural Networks (ANN) and Deep Learning (DL)

- o Theory Behind ANN and DNN
- o Perceptrons for Binary Classification
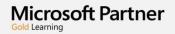- o Getting Started with ANN-binary classification

- Multi-label classification with MLP
- Regression with MLP
- MLP with PCA on a Large Dataset
- Start With Deep Neural Network (DNN)
- Start with H2O
- Default H2O Deep Learning Algorithm
- Specify the Activation Function
- H2O Deep Learning For Predictions

INT

המרכז הבינלאומי
ללימודי הייטק וחדשנות

מתקדמים ✳ **6377**
לקריירה בהייטק

CISCO Networking Academy

Microsoft Partner
Gold Learning

GLOBAL UNIVERSITY SYSTEMS · ibat COLLEGE DUBLIN · HiGHQ מכינים אותך לעולם האמיתי · ARDEN UNIVERSITY · The University of Law · London School of Business & Finance · HTK · TORONTO SCHOOL OF MANAGEMENT

קמפוסים בפריסה ארצית:

באר שבע	**ירושלים**	**רחובות**	**תל אביב**
רחוב האנרגיה 77	רחוב יפו 34	רחוב אופנהיימר 5	ראול ולנברג 36
פארק ההייטק		פארק המדע	קריית עתידים