



DATA DRIVEN BANKING

L'offerta di ABIFormazione
per il 2021

P R O G R A M M I

PERCORSO **PROGRAMMAZIONE PYTHON**

Lo studio del linguaggio Python e l'acquisizione di pratiche di programmazione attraverso esempi concreti ed esercitazioni pratiche.

3 moduli, fruibili anche singolarmente e introduttivi al Percorso professionalizzante "Diventa Data Scientist".



1° MODULO • INTRODUZIONE AI LINGUAGGI DI PROGRAMMAZIONE: IL LINGUAGGIO PYTHON

19, 20 e 21 maggio 2021

► **Introduzione ai linguaggi di compilazione**

- Panoramica linguaggi di programmazione
- Interpreti e Compilatori
- Algoritmi e Computational Thinking

► **Introduzione al Python**

- Rappresentazione dell'Informazione
- Introduzione al linguaggio di programmazione Python
- I notebook Jupyter

► **I tipi di dato Python e costrutti di base**

- Variabili, valori e tipi
- I tipi Interi, Float, Stringhe, Booleani e loro operatori
- Istruzioni condizionali
- Cicli
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Le liste di Python**

- Definizione e Creazione di liste
- Liste annidate
- Iterazione su liste
- Slicing di liste
- Modifica di liste
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Cenni di statistica di base**

- Calcolo e implementazioni di misure statistiche su semplici strutture dati Python

► **Funzioni**

- Definizione di funzioni e scope delle variabili
- Funzioni come parametri
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ 2° MODULO • PROGRAMMAZIONE IN PYTHON

26, 27 e 28 maggio 2021

▶ Processare Liste in Python

- List comprehension
- Ordinamento di liste
- Tipi Mutable e Immutable
- Funzioni anonime
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Stringhe in Python

- Slicing
- Concatenazione
- Attraversamento
- Metodi di manipolazione di stringhe
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Elaborazione di dati strutturati/complessi in Python

- Dizionari e insiemi
- Creazione e accesso
- Interazione di elementi
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Statistica in Python

- Implementazione di funzioni statistiche su tipi eterogenei di dati

▶ 3° MODULO • PROGRAMMAZIONE AVANZATA IN PYTHON

7, 8 e 9 giugno 2021

▶ Elaborazione di serie di dati in Python

- Enumerate e zip
- Map e filter
- Costrutto yield
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Gestione degli errori

▶ Programmazione ad oggetti

- Classi e oggetti
- Attributi e metodi
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Strumenti per la manipolazione dei dati

- Accesso a file di testo, CSV, JSON e Excel
- Pandas
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ Principi di data exploration

- Associazioni e Correlazioni tra variabili
- Distribuzioni e istogrammi
- Regressione lineare
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

PERCORSO PROFESSIONALIZZANTE **DIVENTA DATA SCIENTIST**

Le conoscenze avanzate sulle tecniche e sugli strumenti per la gestione, l'analisi, la visualizzazione di dati e per il loro uso con tecniche di machine learning.

4 moduli, fruibili anche singolarmente, solo chi frequenta l'intero percorso può sostenere il test finale.

Prerequisiti: è richiesta esperienza di programmazione in un qualsiasi linguaggio, in particolare di Python; per chi non ha conoscenze è richiesta la partecipazione al Percorso "Programmazione Python".

1° MODULO • GESTIONE E VISUALIZZAZIONE DEI DATI **16, 17 e 18 giugno 2021**

► **Concetti avanzati per la gestione dei dati in Python**

► **Data Modeling and Understanding: libreria Numpy**

- Tipi di dato Numpy
- Operazioni base su vettori e matrici
- Funzioni universali e broadcasting
- Analisi statistiche avanzate
- I/O con Numpy
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Data Management, Processing and Cleaning: libreria Pandas**

- Serie e Dataframe in Pandas
- Indicizzazione e selezione in Pandas
- Operazioni di base in Pandas
- Metodi per la pulizia dei dati
- Combinazione e aggregazione dei dati
- Integrazione con Microsoft Excel
- Integrazione e omogeneizzazione di diversi data set
- Introduzione alle Espressioni Regolari e pulizia del testo in Pandas
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Data Visualization: librerie Matplotlib, Seaborn, Plotly**

- Istogrammi, fitting, bar graph e tendenze
- Serie temporali, intervalli di confidenza e boxplot
- Scatter plot, heatmap e densità di probabilità
- Aggregazione, selezione e sampling
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

▶ 2° MODULO • TECNICHE DI MACHINE LEARNING PER L'ANALISI DEI DATI 23, 24 e 25 giugno 2021

▶ Introduzione al Machine Learning

- Dati di Training, Validation e Test: definizione e creazione
- Metriche di valutazione dell'efficacia dei modelli di Machine Learning
- Ottimizzazione e selezione del modello prodotto
- Overfitting e Underfitting
- Machine learning con la libreria scikit-learn

▶ Metodi non Supervisionati: Clustering

- Introduzione al problema del Clustering
- Tipologie di Clustering: densità, k-means, gerarchico, ecc.
- Similarità tra oggetti ed ottimizzazione
- Clustering con la libreria scikit-learn
 - K-means
 - Density-based Clustering
 - Hierarchical Clustering
- Hands-on ed esercizi (dati numerici e testo) (Aule virtuali individuali e/o a gruppi)

▶ Metodi Supervisionati: Classificazione

- Introduzione al problema di classificazione
- Classificazione binaria e multi-etichetta
- Matrici di confusione e metriche di bontà della classificazione
- Classificazione con la libreria scikit-learn
 - Logistic Regression, kNN
 - Alberi di Decisione
 - SVM
- Hands-on ed esercizi (dati numerici e testo) (Aule virtuali individuali e/o a gruppi)

▶ Metodi Supervisionati: Regressione

- Introduzione al problema di regressione
- Matrici di confusione e metriche di bontà della classificazione
- Regressione con la libreria scikit-learn
 - Regressione Lineare
 - Ridge, Lasso
 - Alberi di Regressione
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)



3° MODULO • TECNICHE DI DEEP LEARNING PER L'ANALISI DEI DATI 12, 13 e 14 luglio 2021

► **Introduzione al Deep Learning**

- Introduzione alle Reti neurali e al Deep Learning
- Reti neurali multi-livello
- Uso della libreria Keras per l'implementazione di Reti Neurali
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Deep Learning per Classificazione**

- Classificazione Binaria e Multi-Classe
- Definizione della loss di una rete per task di classificazione
- Funzioni di attivazione per task di classificazione
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Deep Learning per Regressione**

- Definizione della loss di una rete per task di regressione
- Funzioni di attivazione per task di regressione
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Deep Learning per Dati Complessi**

- Reti convoluzionali e ricorrenti
- Deep Learning per analisi di serie temporali
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► **Valutazione e Best Practices**

- Valutazione e tuning dei parametri tramite uso del Validation Set
- Overfitting e Underfitting
- Data Augmentation e Dropout per la riduzione dell'overfitting
- Transfer-Learning e generalizzazione
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)



MODULO DI APPROFONDIMENTO



TECNICHE DI TEXT PROCESSING PER L'ANALISI DEI DATI

8, 9 e 10 settembre 2021

► Text Processing in Python

- Pre-processing di testo: pulizia di testo, stopwords, stemming
- Matrice TF-IDF, similarità e classificazione
- Word/Document Embedding (Word2Vec)
- Librerie Python per manipolazione di testo: textblob, NLTK, spacy e gensim
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► Opinion Mining e Sentiment Analysis

- Introduzione al problema e soluzioni allo stato dell'arte
- Derivare opinioni, sentimento e soggettività da testi
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

► Labeling automatico e categorizzazione di testo

- Introduzione al problema e soluzioni allo stato dell'arte
- Tecniche di machine/deep learning classificazione di testo
- Hands-on ed esercizi (Aule virtuali individuali e/o a gruppi)

