

Data Science con Python - dalle stringhe all'intelligenza artificiale - ONLINE

Obiettivi

Il corso intende presentare una panoramica introduttiva sugli strumenti di programmazione offerti da Python per applicazioni di data analysis e machine learning.

Argomenti

Python3 per l'analisi dei dati, dalla loro manipolazione e rappresentazione all'applicazione di tecniche ML, utilizzando una serie di librerie di Python, come pandas, plotly scikit-learn, keras, e con diversi casi d'uso (come: analisi di dati covid, analisi di log, messaggi tweet). Tutte le giornate saranno caratterizzate da 3 ore di teoria accompagnata dallo svolgimento di esercizi ed 1 ora di hands-on.

Prerequisiti

Conoscenza base di python, rudimenti di programmazione per il calcolo scientifico.

Destinatari

Ricercatori, Tecnologi, Tecnici

Date

1 – 4 marzo 2022 (4 giorni) dalle 13.35 alle 18.00

N. partecipanti

20

Responsabile:

Elisabetta Ronchieri (CNAF INFN)

Segreteria organizzativa:

Stefano Dal Pra - stefano.dalpra@cnafe.infn.it

Docenti:

Alessandro Costantini, Stefano Dal Pra, Doina Cristina Duma, Elisabetta Ronchieri

Sede:

online, piattaforma Zoom

PROGRAMMA

I giornata - 1 Marzo

Introduzione Generale – apertura corso;
introduzione al ML, concetti principali;
esempi di infrastruttura per data analysis: (CLOUD @ ML);
markdown language, introduzione a jupyter;
strumenti base di Python per la manipolazione dei dati: pandas, numpy, matplotlib, seaborn, plotly,
dash, dask;
introduzione pandas;
hands-on

II giornata - 2 Marzo

Tecniche ML supervised - strumenti base di Python per ML;
elenco di tecniche ML supervised e misura performances;
Regressione e classificazione;
Classificatori supervised: per esempio Decision Tree, SVM, naive bayes;
NLP: filtro antispam basato su SVM;
hands-on

III giornata - 3 Marzo

Tecniche ML unsupervised - tecniche unsupervised di clustering: per esempio k-means DBMS
Hierarchical clustering;
PCA

IV giornata - 4 Marzo

analisi del testo con NLP, per esempio per analisi di log di un servizio, di tweet;
ANN, CNN, GAN Autoencoder

