



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Big data processing Lab

2122-104R-BDatap-lab

---

#### Obiettivi formativi

Il corso si propone l'obiettivo di introdurre all'utilizzo avanzato di applicazioni data-intensive basate su principi di DataOps, che comprendono principi di orchestrazione di risorse e componenti applicative, infrastrutture software distribuite di tipo cloud hybrid e l'utilizzo di infrastructure code.

#### Contenuti sintetici

In particolare il corso coprirà i seguenti contenuti:

1. Introduzione all'utilizzo di DataOps, Data Warehousing, e principi di Data Quality;
2. Introduzione all'utilizzo di containers e loro orchestrazione;
3. Introduzione all'utilizzo di Serverless computing per data-intensive applications, mediante la suite RADON e altre implementazioni open-source;
4. Introduzione all'utilizzo di approcci MLOps pratici, esemplificate dalla suite DataRobot;

#### Programma esteso

#### Prerequisiti

- Fondamenti di python programming e utilizzo di framework big data (e.g., Apache Spark, Cassandra);

- Fondamenti di architetture software distribuite e loro requirements engineering & design;

## **Metodi didattici**

Il corso si articolerà in lezioni frontali, self-study, invited seminars e interactive sessions.

## **Modalità di verifica dell'apprendimento**

Una sessione pratica di tipo hands-on su una suite nota di tipo MLOps verrà usata per far applicare concretamente i contenuti appresi.

## **Testi di riferimento**

Nessuno (articoli e materiale fornito dal docente).

## **Periodo di erogazione dell'insegnamento**

March 1st - March 4th (virtual).

## **Lingua di insegnamento**

Italian/English.

## **Sustainable Development Goals**

ISTRUZIONE DI QUALITÀ

---