

# DATA SCIENCE FOR THE HUMAN AND SOCIAL SCIENCES (LM81)

(Università degli Studi)

## Teaching Introduction to Data Processing and Machine Learning

GenCod A007245

**Owner professor** ANGELO COLUCCIA

**Teaching in italian** Introduzione all'elaborazione dei dati e Machine

**Teaching** Introduction to Data Processing and Machine Learning

**SSD code** ING-INF/03

**Reference course** DATA SCIENCE FOR THE HUMAN AND SOCIAL SCIENCES

**Course type** Laurea Magistrale

**Credits** 9.0

**Teaching hours** Front activity hours: 54.0

**For enrolled in** 2023/2024

**Taught in** 2023/2024

**Course year** 1

**Language** ITALIAN

**Curriculum** Data Analytics

**Location**

**Semester** First Semester

**Exam type** Oral

**Assessment** Final grade

**Course timetable**

<https://easyroom.unisalento.it/Orario>

### BRIEF COURSE DESCRIPTION

- Introduzione all'elaborazione dei dati e suoi strumenti
  - Elementi base di programmazione, con particolare riferimento alle finalità di elaborazione/processamento digitale dei dati
- Il linguaggio Python
- Introduzione all'intelligenza artificiale e machine learning
- Strumenti matematici e statistici essenziali per il machine learning
- Utilizzo di Python per elaborazione dati e machine learning
- Introduzione alle problematiche sociali ed etiche delle tecnologie ICT e del machine learning

### REQUIREMENTS

Il corso è un corso di base per studenti di area umanistica/sociale, non si assumono prerequisiti tecnici specifici; sono utili conoscenze pregresse di informatica e statistica.

---

## COURSE AIMS

### *Conoscenze e comprensione.*

Al termine del corso gli studenti avranno acquisito le seguenti conoscenze: conoscenza degli elementi della programmazione strutturata e ad oggetti; conoscenza degli aspetti fondamentali dell'acquisizione e elaborazione digitale di dati; conoscenza dei principali modelli di learning, con particolare attenzione a quelli più usati nelle applicazioni di data science; conoscenza dei principali strumenti e approcci di statistical (machine) learning utilizzati in ambito data science.

### *Capacità di applicare conoscenze e comprensione.*

Al termine del corso gli studenti avranno acquisito gli strumenti essenziali per la scrittura di piccoli programmi in Python che eseguano semplici elaborazioni dati e richiamino alcune funzioni principali di interesse per il machine learning. Sapranno definire ed interpretare gli strumenti idonei all'elaborazione di diverse tipologie di dati, in funzione dell'obiettivo applicativo che si intende raggiungere, e individuare quali problemi si prestano ad essere affrontati con approcci di machine learning.

### *Autonomia di giudizio.*

Gli studenti saranno guidati all'apprendimento critico di quanto insegnato in classe, confrontando diversi approcci possibili per uno stesso problema e valutandone pro e contro.

### *Abilità comunicative.*

È essenziale che gli studenti siano in grado di comunicare con interlocutori di vario tipo in modo chiaro, logico ed efficace, utilizzando gli strumenti appresi, le conoscenze e il gergo specialistico. Il corso promuove lo sviluppo di skills quali l'abilità di identificare ed esporre in termini precisi le caratteristiche dei diversi strumenti di elaborazione dei dati e machine learning, nonché l'abilità di descrivere ed analizzare le opzioni disponibili, comunicano in modo rigoroso supportato da argomentazioni tecniche e scientifiche.

### *Capacità di apprendimento.*

Gli studenti dovranno acquisire l'abilità critica di discutere, con originalità ed autonomia, gli aspetti più importanti legati all'elaborazione dei dati ed al machine learning, inclusi aspetti culturali legati alla data science e tecnologie ICT. Essi dovranno essere in grado di sviluppare e applicare le conoscenze acquisite nel prosieguo dei propri studi e in generale nella più ampia prospettiva della propria vita professionale e di lifelong learning.

---

## TEACHING METHODOLOGY

Lezione frontale con ausilio di materiali digitali, utilizzo di strumenti software.

---

## ASSESSMENT TYPE

Esame scritto seguito da eventuale discussione orale.

---

## FULL SYLLABUS

### Introduzione all'elaborazione dei dati

- Glossario minimo di concetti e tecnologie ICT
- Evoluzione tecnologica, hardware e software
- Il paradigma digitale, dall'acquisizione dei segnali alla digitalizzazione
- Rappresentazione delle informazioni, codifica e compressione
- Elaborazione e trasmissione dei dati
  - Uno sguardo all'ecosistema dell'ingegneria dell'informazione (ICT) e alla sua evoluzione: elettronica, automatica/robotica, telecomunicazioni, informatica, cibernetica, algoritmi di machine learning e intelligenza artificiale

### Elementi base di programmazione

- Perché la programmazione
- Tipologie di approcci alla programmazione e linguaggi
- Elaborazione/processamento digitale dei dati
- Introduzione al linguaggio Python
- Variabili e tipi di dati
- Operatori aritmetici e logici
- Costrutti fondamentali
- Funzioni
- Esempi di algoritmi fondamentali
- Debugging

### Introduzione all'intelligenza artificiale e machine learning

- Evoluzione dell'intelligenza artificiale
- La scienza dei dati
- Tipologie di apprendimento automatico
- Il ruolo dei modelli e dell'elaborazione statistica
- Strumenti matematici e statistici essenziali per il machine learning
- Modelli e algoritmi fondamentali per l'apprendimento statistico
- Utilizzo di Python per elaborazione dati e machine learning
- Problematiche sociali ed etiche delle tecnologie ICT e del machine learning

---

## REFERENCE TEXT BOOKS

- Materiale didattico del corso (slides), fornito dal docente
- Allen Downey, "Pensare in Python", Seconda Edizione, Green Tea Press Needham, MA, disponibile gratuitamente [al link https://github.com/AllenDowney/ThinkPythonItalian/blob/master/thinkpython\\_italian.pdf](https://github.com/AllenDowney/ThinkPythonItalian/blob/master/thinkpython_italian.pdf)
- Tutorial ufficiale di Python, <https://pytutorial-it.readthedocs.io/it/python3.10/>
- Chris Mattmann, "Machine Learning with TensorFlow", Second Edition, Manning Publication, Shelter Island, 2020