UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

# COURSE SYLLABUS

# Data Processing and Analysis

**2425-3-E4102B090**

## Aims

The course aims at providing skills (both methodological and technical) and tools for understanding and implementing solutions for Big Data processing (structured and unstructured data), through the use of AI algorithms and tools for the extraction and knowledge representation from real data. In addition, the course intends to provide the technical tools for modelling and realising data models following the NoSQL paradigm, focusing mainly on the graph-database and NoSQL database. Finally, competencies related to XAI will be provided to explain the behaviour of black box algorithms

## Contents

Introduction to AI and Big Data Analytics

Getting knowledge from data

Modelling and Querying the Resulting knowledge

## Detailed program

- **Introduction to AI and Big Data Analytics**

1. Goal and rationale of AI. The relation between Big Data and AI
2. The value of knowledge – digital economy and data-driven decision making

- **Getting knowledge from data**

1. Word Embedding (Word2Vec, Doc2Vec,GLOVE, FastText, StarSpace)
2. Evaluate word embedding models (intrinsic vs extrinsic evaluation)
3. Topic Modelling through Python

**- Modelling and Querying the Resulting knowledge through NoSQL** 1. introduction to NoSQL data stores
2. graph-databases and graph-traversal query languages (Cypher)
3. Document Databases

- **Explainable AI (global and local interpretation models)**

1. Introduction to XAI, local/global interpretation models. model agnostic-specific algorithms
2. XAI techniques as in the state of the art (eg. LIME, SHAP, etc)

## Prerequisites

None

## Teaching form

The course will be provided by means of lessons, seminars, and laboratory sessions and homeworks.

## Textbook and teaching resource

Lectures with the support of slides, laboratory and real-life case studies. Scientific Papers and books indicated by the lecturer. The software used is either available as open source or through academic license

## Semester

III ciclo

## Assessment method

The verification method is based on a written test whilst the oral examination will be provided on request.

The written test takes place at the computer and it consists of open and closed questions with multiple answers on all course topics.

The evaluation is focused on the student's ability to answer to specific questions by referring both to the theoretical and practical aspects (through examples) connected to the requested topic.

The written test is common for both attending students and non-attending students.

The oral exam is aimed at assessing the theoretical knowledge of the student on the topics of the course. The ability to reason and deepen the issues proposed during the examination and the methodological rigor of their development will be evaluated.

## Office hours

By Appointment

## Sustainable Development Goals