UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Laboratorio 2 - Big data analytics: modelli di machine learning e GenAI per l'analisi dei dati

2526-2-F8802N074

Obiettivi formativi

Comprendere le sfide e le opportunità dei Big Data; Acquisire conoscenze su normative europee e strumenti riguardanti la raccolta e l'uso di tracce digitali; Utilizzare modelli di machine learning per l'analisi dati; Integrare l'Intelligenza Artificiale Generativa nella ricerca sociale.

Contenuti sintetici

Definizione ed utilizzi dei Big Data; GDPR e DSA applicati alla raccolta dati digitali nel contesto Europeo; Introduzione al machine learning tramite Weka e Teachable Machine; Utilizzo di modelli linguistici di grandi dimensioni - quali GPT, Claude e Mistral - in diversi compiti della ricerca sociale.

Programma esteso

Lezione 1, 12 Dicembre 2025, 10.30-16.30: Cosa sono e a cosa servono i Big Data
Argomenti principali:
Definizione e caratteristiche dei Big Data (volume, varietà, velocità, veridicità, valore).
Regolamentazioni per l'accesso e l'utilizzo dei Big Data nel contesto europeo: dalla General Data Protection Regulation (GDPR) al Digital Services Act (DSA).
Le tracce digitali come principale fonte di Big Data nella ricerca sociale: fonti e strumenti.
Attività:
Strumenti online "reday-to-use" per i Big Data Analytics.
Esercizi di raccolta di tracce digitali tramite Zeeschuimer

Lezione 2, 17 Dicembre 2025, 10.30-16.30: Introduzione al Machine Learning
Argomenti principali:
Comparazione tra Weka Workbench e Teachable Machine.
Vantaggi e svantaggi dell’open source rispetto a prodotti commerciali chiusi.
Attività:
Esercizi di classificazione dati quantitativi e qualitativi.

Lezione 3, 18 Dicembre 2025, 10.30-16.30: Introduzione ai Modelli Linguistici di Grandi Dimensioni
Argomenti principali:
Comparazione tra GPT, Claude e Mistral.
Vantaggi e svataggi dell’open source rispetto a prodotti commerciali chiusi.
Attività:
Esercizi di trascrizione di interviste con WhisperAI e analisi dei testi tramite Claude.

Lezione 4, 19 Dicembre 2025, 10.30-16.30: Progetto finale
Argomenti principali:
Revisione degli argomenti trattati nel corso del laboratorio.
Implementazione di un progetto finale in piccoli gruppi in stile hackathon.
Attività:
Lavoro di gruppo per la progettazione e realizzazione del progetto finale e presentazione del progetto agli altri partecipanti.

Prerequisiti

Concetti basi di logica della programmazione.

Metodi didattici

Il corso è progettato come un'esperienza pratica di apprendimento, basata sul principio del 'learning by doing'. Il metodo di insegnamento sarà quello della 'didattica capovolta': le letture devono essere preparate in modo autonomo prima delle lezioni, mentre in aula saranno presentati concetti chiave e veranno svolte esercitazioni guidate.

Modalità di verifica dell'apprendimento

La valutazione sarà basata sulla partecipazione attiva durante le lezioni e sulla qualità del progetto finale.

Testi di riferimento

Pilati, F. (2025). Sociologia Digitale 2.0. L’Intelligenza Artificiale Generativa come oggetto e strumento di ricerca. Milano: Ledizioni.

Sustainable Development Goals

ISTRUZIONE DI QUALITÁ