UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Big Data Analytics

2526-3-E3303M016

Obiettivi

Il corso intende fornire le competenze teoriche e pratiche per comprendere metodi e strumenti dedicati all'analisi di grandi moli di dati, con particolare attenzione all'utilizzo di tecniche di data wrangling e machine learning attraverso l'impiego di tool come OpenRefine e Knime. L'insegnamento mira inoltre a sviluppare la capacità di interpretare criticamente i risultati delle analisi, al fine di supportare efficacemente i processi decisionali. Una parte rilevante del corso è dedicata alla comunicazione dei dati e dei risultati analitici mediante tecniche di data visualisation, con l'obiettivo di rendere le informazioni facilmente fruibili e comprensibili. Infine, verranno introdotti i principi fondamentali dell'intelligenza artificiale generativa, con particolare riferimento al suo impiego per la sintesi automatica e la descrizione dei dati.

Contenuti sintetici

Introduzione a Big Data
Data wrangling, machine learning e text mining
Data visualisation

Programma esteso

Introduzione ai Big Data: concetti fondamentali, problematiche di gestione di grandi moli di dati
Data wrangling con OpenRefine: pulizia e trasformazione dei dati
Introduzione al machine learning e text mining
Machine learning e text mining: modelli di apprendimento supervisionato e introduzione al trattamento del testo
Workflow ML base con Knime: costruzione di pipeline di apprendimento automatico, con focus su NLP
Data visualization: tecniche di rappresentazione visiva dei dati e storytelling

AI generativa: sintesi automatica e descrizione dei dati con modelli generativi

Prerequisiti

Nessuno

Modalità didattica

Lezione frontale (DE) in presenza
Laboratori (DI) con strumenti software
Possibilità di 30% ore online sincrone
24h erogative (lezione)
18h interattive (laboratori)
21h esercitazioni + laboratorio

Materiale didattico

Lezioni con l'ausilio di slide, laboratorio e casi applicativi. Articoli scientifici di riferimento saranno forniti dal docente. Il Software utilizzato sarà open-source

Periodo di erogazione dell'insegnamento

Marzo - Aprile

Modalità di verifica del profitto e valutazione

La modalità di verifica si basa su una prova scritta ed una eventuale prova orale.

La prova scritta si svolge al computer ed è composta da domande aperte e chiuse e risposta multipla su tutti gli argomenti del corso.

In sede di valutazione viene considerata la capacità dello studente di rispondere a quesiti specifici facendo riferimento agli aspetti teorici e pratici (mediante esempi) connessi all'argomento richiesto.

La prova scritta è comune sia per gli studenti frequentanti sia per i non frequentanti.

La prova orale è mirata ad accertare la conoscenza teorica dello studente sugli argomenti del corso. Saranno quindi valutate le capacità di ragionare e approfondire le tematiche proposte in sede di esame e il rigore metodologico del loro sviluppo.

Orario di ricevimento

su Appuntamento

Sustainable Development Goals