UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

2627-2-F8205B014-F8205B014-2

Obiettivi formativi

Il laboratorio intende fornire competenze teoriche e, soprattutto, pratiche per comprendere i metodi e gli strumenti dell'intelligenza artificiale applicata all'analisi dei dati, con particolare attenzione alla relazione tra Big Data e AI. Vengono consolidate le basi di programmazione in Python e di interrogazione dei dati con SQL, propedeutiche alle attività di analisi. Una parte rilevante dell'insegnamento è dedicata al text mining e al Natural Language Processing (NLP), dalle rappresentazioni testuali classiche (Bag of Words) fino agli embedding, ai modelli di deep learning e ai Large Language Models (LLM). Il laboratorio introduce inoltre i principi della Explainable AI (XAI), al fine di sviluppare la capacità di interpretare criticamente i modelli e i loro risultati a supporto dei processi decisionali, anche in ambito sanitario ed epidemiologico. Infine, attraverso tecniche di data visualization, si mira a rendere le informazioni e i risultati analitici facilmente fruibili e comprensibili. Il percorso si conclude con la realizzazione di un progetto applicativo di gruppo.

Contenuti sintetici

Introduzione all'AI e ai Big Data. Recap di Python e SQL. Text mining e NLP (Bag of Words, embedding, deep learning, LLM). Explainable AI. Data visualisation.

Programma esteso

Introduzione all'AI e alla Big Data Analytics: obiettivi e razionale dell'intelligenza artificiale; relazione tra Big Data e AI. Recap di Python: elementi fondamentali del linguaggio per l'analisi dei dati. Recap di SQL: teoria delle interrogazioni ed esercitazioni pratiche su basi di dati. Rappresentazione del testo con il modello Bag-of-Words (BoW). Laboratorio di NLP: pre-processing e analisi del testo. Deep Learning ed embedding: reti neurali e rappresentazioni distribuite del testo. Laboratorio di NLP: costruzione di pipeline applicative. Large Language Models (LLM): principi di funzionamento e impieghi. Introduzione alla Explainable AI (XAI): interpretabilità e

spiegabilità dei modelli. Laboratorio di XAI. Laboratorio di Data Visualization: tecniche di rappresentazione visiva e di storytelling dei dati.

Prerequisiti

Conoscenze di base di Machine Learning.

Metodi didattici

Lezione frontale (DE) in presenza. Laboratori (DI) con strumenti software. Possibilità di 30% ore online sincrone. 24h erogative (lezione); 18h interattive (laboratori); 6h esercitazioni + laboratorio.

Modalità di verifica dell'apprendimento

La verifica del profitto si basa sulla realizzazione di un progetto di gruppo, sviluppato in parte durante le ore di laboratorio. Il progetto richiede l'applicazione integrata degli strumenti e dei metodi trattati in laboratorio (analisi dei dati, NLP, modelli di machine e deep learning, XAI e data visualization). In sede di valutazione vengono considerati: la correttezza metodologica e tecnica del lavoro svolto, la capacità di interpretare criticamente i risultati e la qualità della comunicazione dei dati. La valutazione comprende la consegna degli elaborati prodotti e una discussione/presentazione del progetto, durante la quale vengono accertati anche il contributo e la conoscenza teorica dei singoli componenti del gruppo sugli argomenti del corso.

Testi di riferimento

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

Periodo di erogazione dell'insegnamento

Aprile-Giugno

Lingua di insegnamento

Italiano

Sustainable Development Goals

SALUTE E BENESSERE | ISTRUZIONE DI QUALITÁ