UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

2627-2-F8205B015-F8205B015-2

Obiettivi formativi

Il corso introduce i fondamenti teorici e pratici dell'intelligenza artificiale applicata all'analisi di grandi volumi di dati. Partendo da una panoramica sul rapporto tra AI e Big Data, il corso guida gli studenti attraverso le principali tecniche di elaborazione del linguaggio naturale (NLP), apprendimento profondo (Deep Learning) e modelli linguistici di grandi dimensioni (LLM).
L'approccio è fortemente laboratoriale: ogni blocco tematico è affiancato da sessioni pratiche in Python, con l'obiettivo di sviluppare competenze operative direttamente spendibili in contesti reali.

Gli obiettivi del corso sono in linea con i Descrittori di Dublino, coprendo i seguenti aspetti:
• Conoscenza e comprensione
• Capacità di applicare conoscenza e comprensione
• Autonomia di giudizio
• Abilità comunicative
• Capacità di apprendimento

Contenuti sintetici

• Introduzione all'AI e al Big Data Analytics
• Python: riepilogo e strumenti per l'AI
• Text Mining e Natural Language Processing
• Deep Learning
• Word Embeddings
• Transformers
• Large Language Models (LLM) e prompting

Programma esteso

1. Introduzione all'AI e al Big Data Analytics
   • Obiettivi e razionale dell'intelligenza artificiale
   • La relazione tra Big Data e AI: dati come carburante dei modelli
   • Panoramica del percorso del corso
2. Python: riepilogo e strumenti per l'AI
   • Strutture dati, librerie principali (NumPy, Pandas, Scikit-learn)
   • Ambienti di lavoro per il machine learning e il deep learning
3. Text Mining e NLP
   • Bag of Words (BoW) e rappresentazioni vettoriali del testo
   • Lab NLP: preprocessing, tokenizzazione, analisi del testo
4. Deep Learning
   • Architetture di reti neurali: feedforward, RNN, CNN
   • Lab Deep Learning: addestramento e valutazione di modelli
5. Word Embeddings
   • Word2Vec, GloVe e rappresentazioni distribuite del linguaggio
   • Lab NLP Embedding: utilizzo e visualizzazione di embeddings
6. Transformers
   • Architettura Transformer e meccanismo di self-attention
   • Modelli pre-addestrati (BERT, GPT e varianti)
   • Lab NLP Transformers: fine-tuning e applicazioni pratiche
7. Large Language Models (LLM) e prompting
   • Principi di funzionamento degli LLM
   • Tecniche di prompting: zero-shot, few-shot, chain-of-thought
   • Interazione con LLM tramite API (es. OpenAI, Anthropic)
   • Lab NLP LLM: sviluppo di applicazioni basate su LLM
   • Sessioni integrative di laboratorio su casi d'uso reali

Prerequisiti

Conoscenze di base di programmazione Python. È consigliata una familiarità con i concetti fondamentali di statistica e machine learning.

Metodi didattici

Il corso prevede 42 ore di attività suddivise tra lezioni frontali e sessioni di laboratorio in presenza.
Le lezioni teoriche introducono i concetti e le architetture principali; i laboratori in Python consentono di applicare immediatamente quanto appreso su dataset reali. Il corso include inoltre discussioni su casi d'uso concreti dell'AI in ambito aziendale e di ricerca.

Modalità di verifica dell'apprendimento

Le modalità di verifica verranno comunicate dal docente all'inizio del corso.

Testi di riferimento

• Slide delle lezioni e materiale di laboratorio fornito dal docente.
• Articoli scientifici e risorse online indicate durante il corso.
• Il software utilizzato è open-source o liberamente disponibile (Python, Jupyter, librerie HuggingFace, ecc.).

Periodo di erogazione dell'insegnamento

Secondo semestre.

Lingua di insegnamento

Italiano

Sustainable Development Goals

SALUTE E BENESSERE | IMPRESE, INNOVAZIONE E INFRASTRUTTURE