



UNIVERSITÀ DI PAVIA

Anno Accademico 2023/2024

BIG DATA E ALGORITMI DI APPRENDIMENTO AUTOMATICO: CONOSCENZA, INFORMAZIONE, POTERE

Anno immatricolazione	2021/2022
Anno offerta	2023/2024
Normativa	DM270
SSD	SPS/08 (SOCIOLOGIA DEI PROCESSI CULTURALI E COMUNICATIVI)
Dipartimento	DIPARTIMENTO DI SCIENZE POLITICHE E SOCIALI
Corso di studio	COMUNICAZIONE, INNOVAZIONE, MULTIMEDIALITÀ
Curriculum	PERCORSO COMUNE
Anno di corso	3°
Periodo didattico	Primo Semestre (02/10/2023 - 15/12/2023)
Crediti	6
Ore	36 ore di attività frontale
Lingua insegnamento	ITALIANO
Tipo esame	ORALE
Docente	COSTA PAOLO (titolare) - 6 CFU
Prerequisiti	La comprensione degli argomenti svolti nel corso non presuppone specifiche competenze informatiche. La familiarità con Microsoft Excel o altri programmi dedicati alla produzione e alla gestione di fogli elettronici può aiutare. In tal senso si suggerisce agli studenti meno esperti di prepararsi preliminarmente attraverso una delle diverse opportunità di formazione online (es.: Coursera, Udemy, ...) oppure sfruttando gli innumerevoli tutorial disponibili su YouTube. Sono utili anche alcune conoscenze di base di statistica: carattere, distribuzione, media aritmetica, scarto quadratico medio ecc.
Obiettivi formativi	Al termine del corso lo studente deve: a) comprendere i concetti di base relativi ai modelli interpretativi della realtà economica e sociale basati sul paradigma digitale (rappresentazione numerica del mondo e calcolabilità), i big data (le

grandi aggregazioni di dati che fluiscono in tempo reale da molteplici fonti) e le nuove tecniche di intelligenza artificiale (apprendimento automatico e intelligenza artificiale generativa);
b) saper valutare le principali implicazioni sociopolitiche, giuridiche, culturali ed etiche di tali modelli;
c) essere in grado di applicare logiche e tecniche di prompt design nell'uso di alcune piattaforme di intelligenza artificiale generativa (ChatGPT, Midjourney).

Programma e contenuti

Il corso è articolato in tre parti: una parte introduttiva, una di approfondimento e una di laboratorio.

La parte introduttiva (6 lezioni, per un totale di 12 ore) permette di condividere il vocabolario di base necessario alla comprensione dell'oggetto dell'insegnamento: differenza fra dato, informazione e conoscenza; concetto di database e sistema per la gestione di database (DBMS); differenza fra database relazionali e non relazionali; elementi di storia dell'intelligenza artificiale; algoritmi di apprendimento automatico; differenza fra machine learning e deep learning, modelli linguistici (LLMs, Large Language Models) e intelligenza artificiale generativa ecc.

La parte di approfondimento (6 lezioni, per un totale di 12 ore) esplora la natura del fenomeno e le sue radici storiche, evidenziando i fattori che ne determinano la pervasività: esplosione dei big data, "datizzazione" dell'esperienza e diffusione del principio della calcolabilità nei principali ambiti della vita economica e sociale. In particolare, sono presi in considerazione tre domini nei quali il paradigma dei big data sta abilitando i cambiamenti più significativi:

- Informazione e giornalismo
- Marketing e pubblicità
- Bioinformatica, diagnostica medica e ricerca farmaceutica

La parte di laboratorio (6 incontri, per un totale di 12 ore) si organizza intorno all'ambito specifico dell'intelligenza artificiale generativa e delle applicazioni oggi disponibili per l'impiego dei grandi modelli linguistici. In particolare, il laboratorio offre agli studenti la possibilità di addestrarsi nelle principali tecniche di prompt engineering, collaborando con l'intelligenza artificiale per produrre contenuti originali e significativi (testi, immagini, audio e video).

Metodi didattici

Il corso è organizzato in:

- lezioni frontali (accompagnate da verifiche in corsa del livello di apprendimento tramite Wooclap);
- analisi di casi di studio;
- esercitazioni guidate;
- project work di gruppo.

Per lo svolgimento delle lezioni frontali sono utilizzate presentazioni in PowerPoint o PDF, messe a disposizione degli studenti nella sezione dedicata all'insegnamento sulla piattaforma KIRO. Tale materiale è rilasciato in formato conforme ai principi di accessibilità per gli utenti con disabilità (struttura delle intestazioni, ordine di lettura, testo alternativo per le immagini, link autodescrittivi).

La frequenza alle lezioni e alle esercitazioni è fortemente consigliata. Tuttavia, agli studenti che non possono seguire le lezioni viene messa a disposizione la videoregistrazione delle singole lezioni.

Testi di riferimento

Al termine di ogni lezione viene fornita una bibliografia di riferimento relativa agli argomenti trattati. Di seguito si riporta un elenco di monografie utili ad approfondire la tematica del corso nel suo complesso:

- Hannah Fry, Hello World. Essere umani nell'era delle macchine, Bollati Torino, Boringhieri, 2019 (ed. originale Hello World. How to Be Human in the Age of the Machine, London, Black Swan, 2019)
- Astrid Howe, Artificial Intelligence for Beginners: An Introduction to Machine Learning, Neural Networks, and Deep Learning, Astrid Howe, 2023.
- Helga Nowotny, Le macchine di Dio. Gli algoritmi predittivi e l'illusione del controllo, Roma, LUISS University Press, 2022 (ed. originale In AI We Trust. Power, Illusion and Control of Predictive Algorithms, Cambridge, Polity Press, 2021).
- Stuart J. Russell, Peter Norvig, Intelligenza artificiale. Un approccio moderno, Milano, Pearson, 2021 (ed. originale Artificial Intelligence: A Modern Approach. Global Edition, London, Pearson Education, 2021).
- Toby Walsh, Machines Behaving Badly: The Morality of AI, Cheltenham, The History Press, 2022.
- Toby Walsh, Faking It. Artificial Intelligence in a Human World, Cheltenham, The History Press, 2023.

Si precisa che i suddetti testi non sono intesi come letture obbligatorie, da affrontare per la preparazione dell'esame, ma come utili fonti di approfondimento.

Modalità verifica apprendimento

L'esame consiste in una prova orale individuale, della durata di circa 15-20 minuti, volta ad accertare le competenze acquisite relativamente ai contenuti del corso. La prova verte su almeno tre argomenti distinti e relativi ai diversi argomenti trattati. La valutazione finale si basa sul grado di approfondimento e comprensione degli argomenti presentati e sulla capacità di integrare le conoscenze acquisite durante il corso. Il voto finale è assegnato secondo una scala di valutazione da 0 a 30, distribuendo equamente il peso delle singole domande. Durante la prova lo studente può servirsi di materiale multimediale a supporto della sua argomentazione (immagini, filmati, grafici, presentazioni di PowerPoint o simili)

Solo agli studenti che non possono frequentare le lezioni è offerta la possibilità, in alternativa, di sostenere l'esame presentando una relazione orale – eventualmente accompagnata da un supporto multimediale (PowerPoint o simili) – sulla seguente monografia: Luciano Floridi, Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide, Milano, Raffaello Cortina Editore, 2022 (ed. originale rivista, The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities, Oxford, Oxford University Press, 2023).

Altre informazioni

Obiettivi Agenda 2030 per lo sviluppo sostenibile

Nell'ambito dell'insegnamento è prevista la discussione di rischi e opportunità connessi alla diffusione di paradigmi quantitativi e sistemi algoritmici con riferimento agli obiettivi 5 (Parità di genere), 10 (Ridurre le disuguaglianze) e 16 (Pace, giustizia e istituzioni solide).

[\\$lbl legenda sviluppo sostenibile](#)