



UNIVERSITÀ DI PAVIA

Anno Accademico 2022/2023

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

Anno immatricolazione	2020/2021
Anno offerta	2022/2023
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 (26/09/2022 - 16/12/2022)
Crediti	6
Ore	36 ore di attività frontale
Lingua insegnamento	ITALIANO
Tipo esame	ORALE
Docente	COSTA PAOLO (titolare) - 6 CFU
Prerequisiti	<p>Non sono richieste particolari 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.</p>
Obiettivi formativi	<p>L'insegnamento si pone due obiettivi:</p> <ol style="list-style-type: none">1) fornire agli studenti le conoscenze concettuali relative ai modelli interpretativi della realtà economica e sociale basati sul paradigma dei big data (ovvero le grandi aggregazioni di dati che fluiscono in tempo reale da molteplici fonti) e sulle nuove tecniche di intelligenza artificiale

(apprendimento automatico o machine learning);
2) favorire la riflessione sulle implicazioni tecniche, sociopolitiche, giuridiche, culturali ed etiche di tale paradigma.

Programma e contenuti

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

La parte introduttiva (5 lezioni, per un totale di 10 ore) permette di condividere il vocabolario di base necessario alla comprensione del soggetto 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 ecc.

La parte di approfondimento (5 lezioni, per un totale di 10 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 delle logiche algoritmiche 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, medicina e farmaceutica

La parte di laboratorio (5 incontri, per un totale di 10 ore) si organizza intorno all'ambito tematico specifico dell'informazione online. Saranno analizzati specifici casi di studio, con l'intento di discutere con gli studenti:

- l'influenza dei grandi filtri (Google e Facebook) nella formazione dell'opinione pubblica
- le basi tecniche delle cosiddette "fake news" e i principali esempi di canali di disinformazione
- l'evoluzione delle tecnologie per la produzione automatica di notizie
- la pratica e gli strumenti del cosiddetto "data journalism"

Metodi didattici

Il corso è organizzato alternando lezioni frontali ed esercitazioni, con l'obiettivo di familiarizzare con strumenti di fact checking, data scraping e data visualisation.

Testi di riferimento

Si richiede la lettura di uno dei seguenti testi:

- 1) 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", 2018);
- 2) 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).

Modalità verifica apprendimento

La prova di esame si svolgerà in modalità orale. Tutti gli studenti (frequentanti e non frequentanti) potranno liberamente scegliere una delle tre seguenti tre alternative:

- 1) Colloquio della durata 15-20 minuti, finalizzato a verificare la comprensione da parte dello studente dei contenuti del saggio di Hannah Fry "Hello World. Essere umani nell'era delle macchine"

(facoltativi i capitoli su La criminalità e L'arte, obbligatori tutti gli altri);

2) Colloquio della durata 15-20 minuti, finalizzato a verificare la comprensione da parte dello studente dei contenuti del saggio di Helga Nowotny "Le macchine di Dio. Gli algoritmi predittivi e l'illusione del controllo";

3) Presentazione – della durata di 15-20 minuti – di una ricerca personale ("tesina") svolta dallo studente su uno dei temi discussi durante le lezioni, accompagnata da un deck multimediale (PowerPoint o simili). Nel corso dell'esame sarà valutata la capacità di sviluppare un approfondimento in autonomia e di raccogliere informazioni su di esso a partire da una bibliografia autorevole.

Altre informazioni

Obiettivi Agenda 2030 per lo sviluppo sostenibile

Nell'ambito dell'insegnamento si discuteranno 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](#)