

Perché occuparsi di Data Science

Il paradigma Big Data è destinato a diventare uno dei fattori chiave dell'innovazione e della competizione a livello globale. Con il termine Data Science si identifica un campo di studi interdisciplinare, che ha per oggetto i metodi scientifici, le procedure e i sistemi per estrarre conoscenza, comprensione e potenziali previsioni da grandi quantità di dati, sia strutturati che non, nel rispetto della privacy degli individui. Allo scopo, la Data Science impiega teorie e metodi provenienti da vari campi dell'Informatica, della Statistica e della Matematica, ed in particolare da aree quali: Algoritmi, Classificazione, Data Mining, Database, Machine Learning, Metodi Numerici, Ottimizzazione, Sicurezza.

La Data Science fornisce i presupposti scientifici per attività come:

- individuare e campionare sorgenti di dati;
- organizzare e gestire in maniera efficiente grandi quantità di dati, tenendo conto dei vincoli imposti da software, hardware e banda di comunicazione;
- costruire modelli matematici per analizzare regolarità e pattern nascosti nei dati, o anche apprendere da essi;
- garantire che raccolta, trasmissione e analisi dei dati siano condotte senza rischi per la privacy;
- creare visualizzazioni che aiutino la comprensione dei dati;
- presentare e comunicare la conoscenza ricavata dai dati.

Obiettivi formativi

Il Curriculum Data Science mira a fornire una solida preparazione di base sulle tecniche, e i sottostanti principi teorici, che rendono possibile l'analisi dei dati. A tale scopo, nel curriculum vengono combinate e applicate competenze provenienti da diverse aree disciplinari attive nell'Ateneo, principalmente dalle aree di Informatica, Ingegneria dell'Informazione, Matematica e Statistica. In particolare, vengono offerti insegnamenti incentrati sui seguenti aspetti:

- tecniche algoritmiche per l'analisi dei dati, con particolare attenzione a strutture per grandi insiemi di dati e ai

relativi aspetti teorici e pratici;

- algoritmi di data mining per la ricerca di regolarità e pattern nei dati, e strutture dati necessarie alla loro organizzazione;
- metodi crittografici per la protezione della privacy degli individui, durante tutte le fasi di raccolta, trasmissione e analisi dei dati;
- algoritmi di base e avanzati per l'apprendimento statistico, le basi della teoria dell'apprendimento computazionale, il progetto di soluzioni a problemi reali;
- tecniche di programmazione parallela e ad alte prestazioni;
- basi statistiche della regressione, classificazione e inferenza bayesiana, che sono alla base dell'apprendimento automatico;
- metodi numerici per acquisire quelle competenze del computer aided geometric design utili all'implementazione e all'uso di algoritmi specifici per la visualizzazione di dati;
- metodi di ottimizzazione, necessari per condurre efficacemente l'analisi dei dati in presenza di vincoli sulle risorse hardware e software.

Requisiti d'accesso

Per accedere alla Laurea Magistrale in Informatica è necessario:

- aver acquisito un adeguato titolo di studio (laurea in Informatica o in Ingegneria Informatica);
- soddisfare dei requisiti curriculari minimi (avere superato esami equivalenti ad almeno 24 CFU nei settori INF/01 o ING-INF/05 e 24 CFU nei settori MAT/01-09, FIS/01-08 o SECS/01-06);
- avere una adeguata preparazione di ingresso.

Sbocchi professionali

Nella società dell'informazione, quella di Data Scientist sta emergendo naturalmente come una delle professioni più ricercate. Secondo uno studio spesso citato (McKinsey

Global Institute, 2011), la domanda di data scientists nel 2018 potrebbe eccedere di 1,5 milioni la disponibilità effettiva di tali figure. Il laureato magistrale in Data Science sarà in possesso delle competenze per rivolgersi direttamente: ad aziende che, sul territorio o in ambito globale, operano nel campo dell'analisi dei dati di mercato e della *business intelligence*; ad istituzioni che per statuto trattano grandi quantità di dati (medici, finanziari, di censo, etc.); a imprese piccole o grandi che per la gestione delle loro attività si affidano a sistemi informativi complessi.

Il laureato magistrale in Informatica può iscriversi all'Albo degli Ingegneri dell'Informazione (Albo professionale - Sezione A degli Ingegneri - Settore dell'Informazione) e accedere ai dottorati di ricerca in Informatica

Attività didattiche

Insegnamenti I Anno (63 CFU)	CFU	SSD
Data Mining & Organization	12	INF/01
Parallel Computing	6	ING-INF/05
Advanced Algorithms & Graph Mining	6	INF/01
Advanced Statistical Models - A	6	SECS-S/01
Data Security & Privacy	9	INF/01
Corsi a scelta	24	

Insegnamenti II Anno (57 CFU)	CFU	SSD
Bayesian Statistical Inference	9	SECS-S/01
Machine Learning	9	ING-INF/05
Attività di approfondimento	3	
Prova finale: sviluppo lavoro di tesi	21	
Prova finale	3	
Corsi a scelta	12	

I corsi a scelta saranno così distribuiti:

- 12 CFU dovranno essere scelti tra quelli in elenco A
- 6 CFU dovranno essere scelti tra quelli in elenco B
- 6 CFU dovranno essere scelti tra quelli in elenco C
- 12 CFU potranno essere scelti tra quelli in Elenco A, B o C ovvero tra i corsi offerti dall'Ateneo, purché coerenti con il curriculum degli studi. Tra questi corsi possono essere inseriti anche quelli attivati per lauree triennali.