



TOR VERGATA
UNIVERSITÀ DEGLI STUDI DI ROMA

Corso di Laurea in Scienze della Amministrazione e delle Relazioni Internazionali
Anno Accademico 2024/2025

Programma del Corso di Big Data

Modulo Big Data Analytics
CFU 6

Prof. Simone Borra

Contenuti

La multimedialità sempre più pervasiva nella nostra vita e nei nostri processi lavorativi e non, ha portato alla nascita della cosiddetta data-driven economy: un'economia in cui l'uso dei dati è centrale nelle strategie e nella gestione del business, nello sviluppo di nuovi prodotti e servizi, nei processi decisionali. A partire da questo assunto il corso Big Data Analytics si propone di dare allo studente una panoramica sulle tecnologie e le architetture per l'individuazione, l'estrazione, la gestione e l'analisi di grandi quantità di dati. Inoltre, si illustreranno alcune metodologie statistiche per l'analisi multivariata sia nell'ottica della dipendenza sia dell'associazione. L'obiettivo finale è quello di creare nel discente una consapevolezza delle fasi che, a partire da una domanda, portano il professionista a rispondere alla stessa utilizzando le potenzialità che derivano dalla disponibilità di una grande mole di dati.

Obiettivi formativi

Conoscenze e capacità di comprensione (knowledge and understanding)

Gli studenti acquisiscono conoscenze relative alle opportunità di individuare una sorgente di dati derivante da un'economia multimediale. Verranno introdotte le modalità di acquisizione del dato a partire da un problema al quale si vuole dare una risposta. Saranno analizzate le fasi successive di elaborazione di grandi moli di dati, fasi che portano alla scelta di una tecnica di analisi e al dataset statistico associato.

Utilizzazione delle conoscenze e capacità di comprensione (applying knowledge and understanding)

Il corso è progettato per dare il suo contributo alla crescente domanda di laureati che si trovano ad operare in contesti dove è l'informazione ricavata da grandi masse di dati che guida i processi decisionali. Lo studente acquisirà la terminologia adeguata e una panoramica delle tecniche e degli strumenti sottostanti l'analisi dei Big Data. Sarà poi mostrato come utilizzare tali tecniche su dati reali attraverso l'uso di software statistico.

Capacità di trarre conclusioni (making judgements)

Lo studente imparerà a fare le giuste domande per leggere uno studio basato sull'analisi di grandi quantità di dati e capire la robustezza delle conclusioni.

Abilità comunicative (communication skills)

Quello dei Big Data è sempre più un mondo fatto di una sua terminologia, un insieme di strumenti e tecniche, delle modalità specifiche di scrivere report. Lo studente si confronterà con la necessità di

indicare la sorgente del dato, le problematiche di privacy associate, la descrizione delle tecniche di analisi utilizzate, la corretta descrizione dei risultati.

Metodo didattico

Lezioni frontali con slides e con materiale distribuito dal docente per l'analisi di casi di utilizzo basati su studi reali. Si mostrerà l'uso del software statistico SAS per applicare a dati reali le tecniche statistiche studiate.

Contenuti specifici

Tema 1	Big Data: Definizione, le fonti di produzione dei Big Data
Tema 2	Memorizzazione, gestione, interrogazione di grandi masse di dati.
Tema 3	Dai Big Data al dataset statistico : le fasi e le tecniche di raccolta del dato elementare.
Tema 4	Introduzione alla scelta del tipo di modello
	Metodi di apprendimento supervisionato
Tema 5	La regressione lineare multipla
Tema 6	Applicazioni con SAS del modello di regressione lineare
Tema 7	Modello Logit
Tema 8	Applicazioni con SAS del modello Logit
	Metodi di apprendimento non supervisionato
Tema 9	Cluster Analysis
Tema 10	Applicazioni con SAS della Cluster Analysis
Tema 11	Analisi in componenti principali
Tema 12	Applicazioni con SAS dell'Analisi in Componenti Principali

Valutazione

Lo studente dovrà affrontare un esame orale dove gli potrà essere richiesto di svolgere esempi e dimostrazioni. Durante il corso si svolgeranno dei lavori di gruppo facoltativi basati sull'applicazione di metodi statistici multivariati a dataset reali attraverso l'uso del software statistica SAS. La valutazione degli homework di gruppo sarà qualitativa ma servirà alla formulazione del voto finale.

Descrizione delle modalità e dei criteri di verifica dell'apprendimento

La prova di esame valuta, la preparazione complessiva dello studente, la capacità di integrazione delle conoscenze delle diverse parti del programma, la consequenzialità del ragionamento, la capacità analitica e la autonomia di giudizio. Inoltre vengono valutate la proprietà di linguaggio e la chiarezza espositiva, in aderenza con i descrittori di Dublino (1. Conoscenza e capacità di comprensione (*knowledge and understanding*); 2. Capacità di applicare la conoscenza e comprensione (*applying knowledge and understanding*); 3. Autonomia di giudizio (*making judgements*); 4. Capacità di apprendimento (*learning skills*); 5: Abilità di comunicazione (*communication skills*).

Il voto finale sarà rapportato per il 70% al grado di conoscenza e per il 30% alla capacità espressiva (scritte e orali) e di giudizio critico autonomo dimostrate dallo studente.

La prova di esame sarà valutata secondo i seguenti criteri:

Non idoneo: importanti carenze e/o inaccurately nella conoscenza e comprensione degli argomenti; limitate capacità di analisi e sintesi, frequenti generalizzazioni e limitate capacità critiche e di giudizio; gli argomenti sono esposti in modo non coerente e con linguaggio inappropriato.

18-20: conoscenza e comprensione degli argomenti appena sufficiente, con possibili generalizzazioni e imperfezioni; capacità di analisi sintesi e autonomia di giudizio sufficienti, gli argomenti sono esposti in modo frequentemente poco coerente e con un linguaggio poco appropriato/tecnico;

21-23: conoscenza e comprensione degli argomenti di superficie; capacità di analisi e sintesi corrette con argomentazione logica sufficientemente coerente e linguaggio appropriato/tecnico.

24-26: discreta conoscenza e comprensione degli argomenti; buone capacità di analisi e sintesi con argomentazioni espresse in modo rigoroso, ma con un linguaggio non sempre appropriato/tecnico.

27-29: conoscenza e comprensione degli argomenti completa; notevoli capacità di analisi e sintesi. Buona autonomia di giudizio. Argomenti esposti in modo rigoroso e con linguaggio appropriato/tecnico.

30-30L: ottimo livello di conoscenza e comprensione approfondita degli argomenti. Ottime capacità di analisi, di sintesi e di autonomia di giudizio. Argomentazioni espresse in modo originale e con linguaggio tecnico appropriato.

Libri di testo

Materiale didattico a cura del docente.

Orari di ricevimento

Orario e luogo di ricevimento: da definire.

Email docenti: borra@uniroma2.it;

Didattica web

Eventuali pagine di didattica web saranno comunicate dal docente.