UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Big Data in Economics

2627-2-FDS02Q021-FDS02Q02101

Obiettivi formativi

Al termine del corso, lo studente avrà acquisito:

1. Conoscenze e comprensione
   Conoscenze di base e avanzate nell'ambito della data science, dell'informatica e della statistica applicata, con particolare riferimento ai problemi economici e aziendali.
   Comprensione dei concetti fondamentali legati all'inferenza causale, alla previsione e alla classificazione non supervisionata.
2. Capacità di applicare conoscenze e comprensione
   Capacità di utilizzare tecniche di analisi dei dati reali per affrontare problemi concreti in ambito economico e gestionale, anche attraverso casi studio e dataset reali.
   Applicazione di strumenti computazionali e modelli quantitativi a situazioni aziendali complesse.
3. Autonomia di giudizio
   Sviluppo di spirito critico nella scelta degli strumenti metodologici più adatti per ciascun problema (causalità, predizione, clustering).
   Capacità di interpretare risultati statistici in chiave decisionale e strategica.
4. Abilità comunicative
   Capacità di comunicare in modo chiaro ed efficace i risultati dell'analisi, anche con supporto grafico e tecnico, a interlocutori non specialisti in ambito aziendale ed economico.
5. Capacità di apprendimento
   Capacità di approfondire in autonomia strumenti e tecniche di data analysis e machine learning, anche in contesti professionali in evoluzione, grazie a una solida base metodologica.

In generale gli obiettivi formativi ricardono nelle aree di data science, informatica e statistica legate alle imprese

Contenuti sintetici

Il corso è diviso in 4 parti.

Nella prima parte si discute il ruolo dei big data all'interno dell'economia e le nuove sfide da essi generato. Le restanti tre parti del corso discutono sepratamente le tre aree di applicazione (causalità, previsione e classificazione non supervisionata) con esempi concreti riguardanti in modo particolari le funzioni di gestione dei rischi e scelte del consumatore.

Infine saranno organizzati laboratori di utilizzo di algoritmi con il software R.

Programma esteso

1. Introduzione e definizione del problema: the Big Data Challenge

2. Il ruolo dell'incertezza: Causa, previsione e classificazione non supervisionata.

3. Meccanismi causali: elementi fondamentali e un case study.

4. Previsione: la sfida di valutare l'incertezza nei modelli predittivi.

5. Apprendimento non supervisionato: Self-Organizing-Map e marketing

6. Bonus track: l'analisi del Drift nei modelli aziendali.

7. Reporting dei risultati in impresa: creare una narrativa intorno ai modelli.

Prerequisiti

Principi di programmazione in R o pyhton. Statistica di base

Metodi didattici

Il corso di 21 ore è composta da

11 ore di lezioni laboraratorio in remoto modalità sincrona.
10 ore di lezioni in presenza con modalità erogativa.

Modalità di verifica dell'apprendimento

Esame scritto e progetto di analisi dati (50%-50%) . Entrambi verteranno sugli argomenti del corso. Durante l'esame sarà valutata la correttezza e la chiarezza delle risposte.

Testi di riferimento

Il materiale didattico è costituito da alcuni capitoli di libro e da una serie di articoli su rivista scientifica e dai lucidi delle lezioni. Il materiale didattico sarà disponibile in forma elettronica sulla piattaforma di e-learning. Quando questo non è possibile una copia cartacea degli articoli sarà resa disponibile all'inizio del corso.

libro:

Data Science for Business
What You Need to Know about Data Mining and Data-Analytic Thinking
By Foster Provost, Tom Fawcett

Periodo di erogazione dell'insegnamento

II semestre

Lingua di insegnamento

Inglese

Sustainable Development Goals

IMPRESE, INNOVAZIONE E INFRASTRUTTURE