



**UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA**

SYLLABUS DEL CORSO

Laboratory of Data Management and Synthetic Indicators for Sustainability

2627-2-F7603Q028

Obiettivi

Al termine del corso, gli studenti saranno in grado di definire, implementare e interpretare indicatori sintetici per misurare la sostenibilità e di applicare i principali approcci di machine learning per l'analisi di sistemi di dati complessi definiti da macro-indicatori di sostenibilità. In particolare, saranno in grado di:

- definire indicatori sintetici e padroneggiare gli elementi fondamentali dell'analisi multivariata e del processo decisionale multicriterio, selezionando il metodo più appropriato (conoscenza e capacità di comprensione)
- illustrare i metodi, interpretare criticamente gli indicatori costruiti e valutare le caratteristiche, i campi di applicazione, i vantaggi e gli svantaggi delle singole tecniche di machine learning (conoscenza e capacità di comprensione; autonomia di giudizio)
- condurre analisi statistiche utilizzando software statistici per implementare e calcolare gli indicatori, scegliendo e gestendo gli approcci di indagine più adatti alle finalità dell'analisi (capacità di applicare conoscenza e comprensione)
- comunicare i risultati e illustrare i metodi a un pubblico che potrebbe non avere familiarità con i metodi statistici o di machine learning (abilità comunicative)
- apprendere in modo autonomo ed esplorare ulteriori estensioni metodologiche o nuovi indicatori nazionali e internazionali (capacità di apprendimento)

Contenuti sintetici

- Definizioni di modello concettuale e di indicatore
- Costruzione degli indicatori
- Calcolo mediante software statistico e interpretazione degli indicatori
- Esempi di indicatori nazionali e internazionali per la sostenibilità
- Struttura e pre-trattamento dei dati nell'ambito del machine learning; covarianza e correlazione

- Strategie per l'esplorazione dei dati e la razionalizzazione di problemi complessi: Analisi delle Componenti Principali
- Misure di distanza per la similarità e la diversità
- Modellazione supervisionata per l'analisi degli indicatori di sostenibilità
- Strategie per il processo decisionale multicriterio
- Esperienze pratiche e casi di studio

Programma esteso

Definizione del modello concettuale e dell'indicatore:

- Introduzione alle definizioni fondamentali dei quadri concettuali e degli indicatori
- Classificazione delle tipologie di indicatori in base al fenomeno analizzato, all'ambito, al contenuto informativo e alla struttura di costruzione

Fasi per la costruzione degli indicatori:

- Sviluppo di un solido quadro teorico
- Criteri per la selezione delle variabili e screening dei dati
- Trattamento dei dati mancanti e tecniche di imputazione
- Metodi di normalizzazione, ponderazione e aggregazione

Calcolo mediante software statistico e interpretazione degli indicatori:

- Implementazione computazionale pratica degli indicatori mediante software statistico
- Linee guida pratiche per la corretta interpretazione degli indici sintetici

Esempi di indicatori nazionali e internazionali per la sostenibilità:

- Rassegna dei quadri internazionali (ad es. indicatori progettati per monitorare i progressi degli Obiettivi di Sviluppo Sostenibile dell'ONU)

La struttura dei dati multivariati:

- Covarianza e correlazione
- Scalatura e pre-trattamento dei dati: centratura, auto-scaling, range scaling, variance scaling

Strategie per l'esplorazione dei dati e la razionalizzazione di problemi complessi:

- Analisi della struttura ed esplorazione dei dati relativi a sistemi complessi nell'ambito dei macro-indicatori sintetici per la sostenibilità
- Analisi delle Componenti Principali (PCA): scores e loadings; autovalori e definizione delle componenti significative (analisi del rango)
- Esempi di PCA e casi di studio

Analisi della similarità e della diversità nei sistemi complessi:

- I concetti di analogia, similarità, dissimilarità e distanza
- Misure di distanza e similarità per dati quantitativi e binari
- Cluster Analysis: esempi e casi di studio

Modellazione supervisionata per l'analisi degli indicatori di sostenibilità:

- Il concetto di bias e i metodi di validazione; tecniche di validazione dei modelli supervisionati
- Metodi di regressione e classificazione multivariata: esempi e casi di studio

Strategie per il processo decisionale multicriterio:

- Funzioni di desiderabilità e di utilità per l'ordinamento razionale di dati multivariati; fronte di Pareto
- Esempi e casi di studio

Prerequisiti

Principi della statistica descrittiva

Modalità didattica

Synthetic Indicators for Sustainability

Lezioni frontali e attività pratiche di laboratorio:

- 10 ore, lezioni in presenza, didattica erogativa
- 10 ore, lezioni in presenza, didattica interattiva
- 10 ore, lezioni in presenza, didattica interattiva per la realizzazione di una relazione su un caso di studio, da analizzare a gruppi da parte degli studenti

Machine Learning and Ai-Applications for Data Management

Lezioni frontali e attività pratiche di laboratorio:

- 10 ore, lezioni in presenza, didattica erogativa
- 10 ore, lezioni in presenza, didattica interattiva
- 10 ore, lezioni in presenza, didattica interattiva per la realizzazione di una relazione su un caso di studio, da analizzare a gruppi da parte degli studenti

Materiale didattico

- OECD/European Union/EC-JRC (2008), Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Publishing, Paris, <https://doi.org/10.1787/9789264043466-en>.
- Mazziotta, M., & Pareto, A. (2025). Statistics for Composite Indicators. Springer.
- Slides
- Articoli scientifici resi disponibili sul sito e-learning del corso
- Dati e codice per le esercitazioni pratiche resi disponibili sul sito e-learning del corso

Periodo di erogazione dell'insegnamento

Primo semestre

Modalità di verifica del profitto e valutazione

Gli studenti dovranno preparare un report e una presentazione su un caso di studio fornito dai docenti e sviluppato in gruppo. La presentazione si terrà durante l'ultima lezione del corso.

L'esame consiste in una prova orale, che comprende la discussione degli argomenti trattati nel corso e la discussione del report, con una votazione compresa tra 18 e 30/30. La presentazione costituisce parte integrante della valutazione: aggiunge fino a un massimo di 2 punti al voto della prova orale, solo se quest'ultima ha riportato un voto maggiore o uguale a 18.

La valutazione si concentra sui seguenti punti: (1) conoscenza e capacità di comprensione; (2) uso corretto della terminologia; (3) capacità di analizzare i dati e valutare criticamente i risultati; (4) capacità di spiegare i metodi utilizzati e comunicare i risultati.

Orario di ricevimento

Su appuntamento via email

Sustainable Development Goals

ISTRUZIONE DI QUALITÀ
