UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Advanced Data Management and Decision Support Systems

2627-1-F9103Q008

Obiettivi

Modulo Advanced Data Management
L'obiettivo principale del modulo "Advanced Data Management" è quello di introdurre i principali strumenti concettuali, teorici e pratici atti alla gestione del dato. Il modulo si pone, in particolare, quattro obiettivi: 1) Fornire agli studenti le basi pratico-teoriche dei sistemi di data management, inclusi temi legati alla distribuzione del dato; 2) Illustrate e introdurre i principali paradigmi di rappresentazione, memorizzazione e manipolazione del dato; 3) Fornire una solida base ai concetti fondamentali al processamento del dato in contesti large-scale; 4) Fornire un'introduzione ai principali concetti di data governance, data quality e data engineering.

Alla fine del corso, gli studenti dovrebbero acquisito ed essere in grado di dimostrare la loro conoscenza nei seguenti argomenti (DdD1, DdD4):

- Basi e principi di design dei sistemi di data management
- Modelli concettuali, logici e fisici per la rappresentazione e gestione del dato
- Principi di data management in contesti distribuiti (replicazione, sharding, consenso distribuito, transazioni)
- Fondamenti di data processing large-scale (MapReduce/Hadoop, Spark)
- Principi di data governance, data quality e data engineering

Dovrebbero inoltre aver acquisito l'abilità di applicare le nozioni impartite in contesti pratici, in particolare (DdD2):

- Abilità di tradurre requisiti utente in un modello dei dati e sua implementazione
- Abilità di gestire dati utilizzando diversi linguaggi di query (SQL, Cipher)
- Abilità di implementare pipeline di gestione e processamento del dato in contesti large-scale, utilizzando strumenti come Apache Spark

Dovrebbero infine aver raggiunto un grado di maturità sufficiente a permettere di comprendere autonomamente gli sviluppi nel settore del data management (DdD5), incluse la capacità di valutare criticamente diverse soluzioni di gestione del dato, loro benefici e limitazioni (DdD3).

Modulo Decision Support Systems
L'obiettivo del corso è duplice: da un lato, introdurre gli studenti ai principali strumenti concettuali e teorici per modellare scelte razionali nel decision-making; dall'altro lato, fornire agli studenti modelli e strumenti che consentano di progettare sistemi di supporto alle decisioni usabili (i.e., efficaci, efficienti e facili da usare) e di valutarli nel mondo reale.

Alla fine del corso, gli studenti dovrebbero aver acquisito ed essere in grado di dimostrare la loro conoscenza nei summenzionati ambiti, in particolare (DdD1, DdD4):
- Modelli e definizioni del decision-making
- Fondamenti teorici del decision-making
- Automazione di processi decisionali
- Dimensioni di qualità dei sistemi di supporto alle decisioni: usabilità, fiducia, dipendenza, compliance, reliance

Dovrebbero inoltre aver acquisito l'abilità di applicare le nozioni acquisite in contesti pratici, in particolare (DdD2):
- Abilità di definire e analizzare scenari di decision-making
- Abilità di progettare e valutare sistemi di supporto alle decisioni

Dovrebbero infine aver raggiunto un grado di maturità sufficiente a permettere di comprendere autonomamente gli sviluppi nel settore della scienza delle decisioni (DdD5), incluse la capacità di valutare criticamente diverse soluzioni di supporto alla decisione, loro benefici e limitazioni (DdD3).

Contenuti sintetici

Modulo Data Management
Basi di data management moderno
Modelli del dato e query relazionali e NoSQL
Distributed data management
Distributed data processing

Modulo Decision Support Systems
Sistemi a supporto delle decisioni: definizioni e principi di dessign
Elementi di teoria delle decisioni: singolo agente, multi-agente, sequenziale
Elementi di teoria delle decisioni comportamentale: razionalità limitata, euristiche, bias
Principi di valutazione di DSS

Programma esteso

Modulo Data Management

Fondamenti dei sistemi di gestione dei dati (lezione erogativa, 1h)
Principi di progettazione: affidabilità, scalabilità, manutenibilità (lezione erogativa, 1h)
Modelli di dati (relazionale, key-value, documentale, a grafo) e linguaggi di interrogazione (lezione erogativa, 2h)
* Modello concettuale dei dati (modello Entity-Relationship)
* Modelli logici dei dati (relazionale, key-value, documentale, a grafo, embedding vettoriali)
* Modelli fisici dei dati e formati di memorizzazione

Utilizzo pratico di DBMS relazionali e SQL (lezione laboratoriale, 4h)
Utilizzo pratico di DBMS NoSQL (Redis, MongoDB, Neo4J) (lezione laboratoriale, 4h)
Transazioni e proprietà ACID (lezione erogativa, 2h)
Gestione distribuita dei dati (replicazione, sharding, consenso) (lezione erogativa, 4h)
Elaborazione distribuita dei dati (MapReduce, Hadoop, Spark) (lezione erogativa, 2h)
Utilizzo pratico di Apache Spark (lezione laboratoriale, 4h)
Principi di Data Governance, sicurezza dei dati e ingegneria responsabile dei dati (lezione erogativa, 4h)

Modulo Decision Support Systems
Tutte le lezioni sono in forma erogativa.

- Fondamenti dei Decision Support Systems (4h)
    - Definizioni di Decision Support Systems
    - Architetture dei DSS
    - Principi di progettazione dei DSS
- Teoria della decisione e teoria dell'utilità attesa (4h)
    - Teoria della decisione per agente singolo (decisione in condizioni di ignoranza, decisione in condizioni di rischio)
    - Aspetti computazionali del decision-making (reti bayesiane, reti decisionali)
- Teoria dei giochi non cooperativi (4h)
    - Concetti di soluzione (ottimalità paretiana, equilibri di Nash, soluzioni minimax/maximin, equilibri correlati)
    - Aspetti computazionali della teoria dei giochi (algoritmi e complessità computazionale)
    - Fondamenti di mechanism design
- Decision-making sequenziale e Reinforcement Learning (4h)
    - Processi decisionali di Markov e programmazione dinamica
    - Fondamenti di Reinforcement Learning (trade-off esplorazione/sfruttamento, Q-learning)
    - Ricerca e pianificazione (Monte Carlo Tree Search)
- Fondamenti di approcci alternativi ai DSS (4h)
    - DSS basati sulla conoscenza (basi di conoscenza, motori inferenziali)
    - DSS conversazionali e basati su IA generativa
- Teoria comportamentale della decisione, euristiche e bias (4h)
    - Razionalità limitata
    - Euristiche
    - Bias cognitivi
    - Razionalità ecologica e decision-making naturalistico
    - Prospect Theory

- Valutazione dei Decision Support Systems (4h)
    - Principi di valutazione quantitativa (metriche di accuratezza e qualità decisionale, progettazione sperimentale)
    - Oltre le metriche quantitative (robustezza, fairness, explainability)
    - Qualità dell'interazione uomo-DSS

Prerequisiti

Modulo Data Management
Conoscenza di base del modello relazionale e del linguaggio SQL

Modulo Decision Support Systems
Nozioni di basi di teoria delle probabilità, intelligenza artificiale e sufficiente maturità matematica

Modalità didattica

Modulo Data Management
Lezioni erogative frontali (16h), esercitazioni di laboratorio computer-based (12h).
Tutte le lezioni vengono svolte in presenza.

Module Decision Support Systems
Lezioni erogative frontali (28h)
Tutte le lezioni vengono svolte in presenza.

Materiale didattico

Module Data Management
Slide fornite dal docente.

Libri di testo (consigliati)
Kleppmann, Martin. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems. Sebastopol, CA: O'Reilly Media, 2017. ISBN: 978-1449373320.
Silberschatz, Abraham, Henry F. Korth, and S. Sudarshan. Database System Concepts. 7th ed. New York: McGraw-Hill Education, 2019. ISBN: 978-0078022159.
van Steen, Maarten, and Andrew S. Tanenbaum. Distributed Systems. 4th ed. CreateSpace Independent Publishing Platform, 2023. ISBN: 978-1718502642.
Karau, Holden, Andy Konwinski, Patrick Wendell, and Matei Zaharia. Learning Spark: Lightning-Fast Data Analytics. 2nd ed. Sebastopol, CA: O'Reilly Media, 2020. ISBN: 978-1492050049.
Redmond, Eric, and Jim R. Wilson. Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement. 2nd ed. Raleigh, NC: Pragmatic Bookshelf, 2018. ISBN: 978-1680502534.

Eventuale materiale aggiuntivo sarà reso disponibile tramite e-learning.

Modulo Decision Support Systems
Slide fornite dal docente.

Libri di testo (consigliati)
Peterson, Martin. An Introduction to Decision Theory. 2nd ed. Cambridge: Cambridge University Press, 2017. ISBN: 978-1107151598.
Shoham, Yoav, and Kevin Leyton-Brown. Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. Cambridge: Cambridge University Press, 2008. ISBN: 978-0521899437.
Kochenderfer, Mykel J., Tim A. Wheeler, and Kyle H. Wray. Algorithms for Decision Making. Cambridge, MA: MIT Press, 2022. ISBN: 978-0262047012.
Wickens, Christopher D., Justin G. Hollands, Simon Banbury, and Raja Parasuraman. Engineering Psychology and Human Performance. 5th ed. New York: Routledge, 2021. ISBN: 978-1138087514.
Gigerenzer, Gerd. Rationality for Mortals: How People Cope with Uncertainty. Oxford: Oxford University Press, 2008. ISBN: 978-0195171143.
Russell, Stuart. Human Compatible: Artificial Intelligence and the Problem of Control. New York: Viking, 2019. ISBN: 978-0525558613.

Eventuale materiale aggiuntivo sarà reso disponibile tramite e-learning.

Periodo di erogazione dell'insegnamento

Secondo semestre

Modalità di verifica del profitto e valutazione

Modulo Data Management
Il modulo prevede due diverse (esclusive) modalità d'esame, a scelta dello studente (in accordo col docente):

Esame scritto (domande a risposta aperta) sui contenuti del corso.
Gli studenti dovranno rispondere a 5 domande aperte che possono vertere su tutti i contenuti del corso, in cui verrà richiesta una discussione critica circa diversi aspetti dei sistemi di data managemente. Ogni domanda vale fino a 6 punti, valutati sulla base dei seguenti elementi: completezza, correttezza, comunicazione e proprietà linguaggio.

OPPURE

Progetto implementativo relativo agli argomenti del modulo E esame orale sui contenuti del corso.
Il progetto può essere svolto sia individualmente che in piccoli gruppi (2/3 studenti), l'esame orale è individuale.
Il progetto verrà valutato fino a 24 punti sulla base dei seguenti elementi: completezza, correttezza, comunicazione e reporting.
L'esame orale consisterà di tre domande sui contenuti del corso, ciascuna delle quali verrà valutata fino a 3 punti.

Il modulo non prevede prove in itinere.

Modulo Decision Support Systems
Esame scritto (domande aperte) sui contenuti del corso (voto massimo 30). Gli studenti dovranno rispondere a 5 domande aperte che possono vertere su tutti i contenuti del corso, in cui verrà richiesta una discussione critica circa diversi aspetti dei sistemi di supporto alle decisioni. Ogni domanda vale fino a 6 punti, valutati sulla base dei seguenti elementi: completezza, correttezza, comunicazione e proprietà linguaggio.

Il corso prevede inoltre la possibilità di svolgere un progetto originale opzionale (saggio scritto) per gli studenti il cui voto nello scritto >= 18 (massimo 4 punti aggiuntivi).

Il modulo non provede prove in itinere.

Orario di ricevimento

Modulo Data Management
Su appuntamento (email: andrea.campagner@unimib.it).
Module Decision Support Systems
Su appuntamento (email: andrea.campagner@unimib.it).

Sustainable Development Goals

ISTRUZIONE DI QUALITÁ