



DENOMINAZIONE INSEGNAMENTO: INFORMATICA		
SSD:INF\01	CFU 6 (3+3)	ORE 18+18
DOCENTE TITOLARE DELL'INSEGNAMENTO: ANTONIO CELESTI		
EMAIL: acelesti@unime.it		
OBIETTIVI FORMATIVI		
Modulo 1		
<p>Comprendere il funzionamento dei sistemi informativi. Conoscere la teoria sulla quale si fondano i principali DataBase Management System (DBMS) con particolare riferimento al modello e all'algebra relazionale. Conoscere la teoria sulla quale si fonda l'interrogazione nelle basi di dati. Comprendere l'architettura delle banche dati web. Conoscere l'evoluzione dei database, da quelli centralizzati per i piccoli e medi sistemi aziendali a quelli paralleli e distribuiti per la gestione dei Big Data nel Cloud. Saper consultare con efficacia le banche dati web, con particolare riferimento a quelle dedicate alla ricerca scientifica.</p>		
Modulo 2		
<p>Acquisire competenze riguardanti la progettazione, lo sviluppo e la gestione di database relazionali. Fornire una panoramica molto ampia sulle metodologie, le tecnologie e i principali strumenti di data science con particolare riferimento al data analytics, data warehousing, data mining e all'intelligenza artificiale. Fornire le competenze necessarie per la progettazione e lo sviluppo di software avanzati, capaci di manipolare ed elaborare dati mediante tecniche di machine learning atte a predire risultati e ad estrarre conoscenza.</p>		
CONTENUTI DEL CORSO		
Modulo 1		
1. I Sistemi informativi		
<p>informazioni e dati; basi di dati e sistemi di gestione di basi di dati; modelli dei dati; schemi e istanze; livelli di astrazione nei DBMS; indipendenza dei dati; linguaggi e utenti delle basi di dati; vantaggi e svantaggi dei DBMS.</p>		
2. Il modello relazionale.		
<p>Modelli logici; relazioni e basi di dati; vincoli di integrità e di tupla; chiavi e valori nulli; vincoli di integrità referenziale.</p>		



3. Algebra relazionale

Unione, intersezione, differenza; Ridenominazione; selezione; proiezione; join; interrogazioni in algebra relazionale; equivalenza di espressioni algebriche; viste; calcolo relazionale; datalog.

4. Architetture delle Banche Dati nel Web.

Introduzione; tipi di architetture; definizione dei dati in SQL; i domini elementari; definizione di schema; definizione di domini e tabelle; vincoli intra-relazionali; vincoli inter-relazionali; modifica di schemi; cataloghi relazionali; interrogazioni in SQL; dichiaratività di SQL; interrogazioni semplici; operatori aggregati; interrogazioni con raggruppamento; interrogazioni di tipo insiemistico; interrogazioni nidificate; modifica dei dati in SQL: inserimento, cancellazione, modifica.

5. Big Data e Basi di Dati per il Cloud

Big Data: caratteristiche; teorema CAP; Database paralleli e distribuiti; DBMS NoSQL; lo storage dei dati nel Cloud.

6. Consultazione Efficace di Banche Dati della Ricerca

Bibliometria; indici di impatto di rivista e di pubblicazione; valutazione della qualità ricerca nel sistema universitario Italiano e internazionale; prestigio editoriale; banche dati bibliometriche e non bibliometriche; esempi con banche dati: Scopus, WoS e Pubmed.

Modulo 2

1. Amministrazione di DBMS Relazionali

L'ambiente MySQL-PHPmyAdmin, creazione, modifica, cancellazione di database, tabelle, righe, colonne e vincoli di chiave; Esecuzione di query.

2. L'Analisi dei Dati

Introduzione ai Big Data e data science; il ciclo di vita dei dati; elaborazione, trasformazione, integrazione e produzione dati; data warehouse: definizione, OLAP: tipologie, operazioni roll-up, drill-down, slice & dice e pivotting; business intelligence: definizione; finalità, strumenti; data mining; definizione; finalità, strumenti; data analytics;

3. Machine Learning



Evoluzione dei sistemi di analisi dati; tecniche supervisionate, non supervisionate, semi-supervisionate, per rinforzo; algoritmi di classificazione, regressione e clustering; Introduzione al deep learning.

4. Introduzione al Linguaggio R

finalità, vantaggi e svantaggi; la console interattiva; l'ambiente R-Studio; esecuzione di programmi; le variabili; tipi di dato; funzioni; strutture dati: vettori, matrici, array, liste, data frame; operatori logici e strutture di controllo; funzioni per la manipolazione dei dati; i packages; gli script.

5. R Avanzato per il Machine Learning

Importazione dati da file e DBMS; Comandi per la grafica 2D e 3D; Funzioni di confronto tra distribuzioni dati; Funzioni per la gestione di Big Data; Approccio pratico: definizione del problema, preparazione del dataset, valutazione di algoritmi; miglioramento e presentazione dei risultati.

TESTI DI RIFERIMENTO

- Atzeni, Ceri, Paraboschi, Atzeni, Ceri, Paraboschi, Torlone Torlone. Basi di dati McGraw-Hill
- Alessandro Rezzani. "Big data analytics". Apogeo Education.
- Sinan Ozdemir. "Data Science". Apogeo Education.
- J. Chambers Software for Data Analysis: Programming with R, Springer.

RISORSE WEB

- <https://www.anvur.it/attivita/vqr/>
- <https://www.scopus.com/>
- <https://apps.webofknowledge.com/>
- <https://pubmed.ncbi.nlm.nih.gov/>
- <https://cran.r-project.org/>
- <http://cran.r-project.org/doc/manuals/R-intro.pdf>
- <https://rstudio.com/products/rstudio/download/#download>