SCUOLA	Scienze di Base ed Applicate
ANNO ACCADEMICO	2014/2015
CORSO DI LAUREA (o LAUREA MAGISTRALE)	LM Informatica
INSEGNAMENTO	Big Data Management
TIPO DI ATTIVITÀ	Caratterizzante
AMBITO DISCIPLINARE	Discipline Informatiche
CODICE INSEGNAMENTO	
ARTICOLAZIONE IN MODULI	NO
SETTORI SCIENTIFICO DISCIPLINARI	INF/01
DOCENTE RESPONSABILE	Simona Ester Rombo Ricercatore Universitario Università degli Studi di Palermo
CFU	6
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	102
NUMERO DI ORE RISERVATE ALLE ATTIVITÀ DIDATTICHE ASSISTITE	48
PROPEDEUTICITÀ	Nessuna
ANNO DI CORSO	Primo
SEDE DI SVOLGIMENTO DELLE LEZIONI	Dipartimento di Matematica ed Informatica- Vedere calendario sul sito www.cs.unipa.it
ORGANIZZAZIONE DELLA DIDATTICA	Lezioni frontali in aula e in laboratorio
MODALITÀ DI FREQUENZA	Facoltativa
METODI DI VALUTAZIONE	Prova progettuale da realizzare in gruppo, prova orale
TIPO DI VALUTAZIONE	Voto in trentesimi
PERIODO DELLE LEZIONI	Secondo semestre
CALENDARIO DELLE ATTIVITÀ DIDATTICHE	Vedere il Calendario del Corso di Laurea sul sito www.cs.unipa.it
ORARIO DI RICEVIMENTO DEGLI STUDENTI	Da concordare via e-mail

RISULTATI DI APPRENDIMENTO ATTESI

Conoscenza e capacità di comprensione

Acquisizione di conoscenze sull'organizzazione dei sistemi informativi aziendali, sull'utilizzo e la progettazione dei sistemi di supporto alle decisioni, sulle problematiche correlate al big data management.

Capacità di applicare conoscenza e comprensione

Capacità di analizzare problematiche relative a contesti caratterizzati da grosse moli di dati e proporre soluzioni progettuali per la gestione e analisi dei dati in tali contesti ed il supporto alle decisioni.

Autonomia di giudizio

Capacità di analizzare e valutare soluzioni per la gestione di grosse quantità di dati. Capacità di progettare sistemi di supporto alle decisioni mediante l'analisi delle specifiche tecniche fornite.

Abilità comunicative

Capacità di descrivere soluzioni di progettazione di sistemi informativi complessi e sistemi di supporto alle decisioni e analizzarne le prestazioni. Capacità di cooperare per determinare delle soluzioni di progettazione appropriate in contesti applicativi caratterizzati dalla presenza di grosse quantità di dati.

Capacità d'apprendimento

Capacità di aggiornamento attraverso la consultazione di testi avanzati e di pubblicazioni scientifiche correlate agli argomenti affrontati durante il corso. Capacità di seguire, utilizzando le conoscenze acquisite, sia master di secondo livello, che corsi di approfondimento e seminari specialistici nell'ambito della Business Intelligence.

OBIETTIVI FORMATIVI DEL MODULO

Obiettivo del corso di Big Data Management è fornire agli studenti di Informatica conoscenze sull'organizzazione dei sistemi informativi aziendali, sull'utilizzo e la progettazione dei sistemi di supporto alle decisioni, sulle problematiche correlate alla gestione di grosse quantità di dati. Il corso inizierà con una panoramica su Business Intelligence e si concentrerà su progettazione di Data Warehouse e OLAP, su cui verteranno anche gran parte delle attività di laboratorio. Il corso fornirà anche nozioni relative alla progettazione di database non relazionali che si stanno recentemente diffondendo nel contesto dei big data e mostrerà degli scenari applicativi reali nell'ambito di reti sociali e dati biologici.

ORE FRONTALI	LEZIONI FRONTALI
2	BUSINESS INTELLIGENCE E SISTEMI INFORMATIVI AZIENDALI Acquisizione dei dati. Elaborazione dei dati al fine di produrre informazioni. Archiviazione dei dati grezzi e delle informazioni. Trasmissione dei dati e delle informazioni. Presentazione dei dati e delle informazioni.
6	DATA WAREHOUSING Sistemi di supporto alle decisioni. Architetture per il Data Warehousing: architettura ad un livello, architettura a due livelli, architettura a tre livelli. I metadati. Qualità di un Data Warehouse. Il modello multidimensionale e l'OLAP. Le principali operazioni OLAP. Modelli logici a supporto del Data Warehousing.
6	PROGETTAZIONE DI UN DATA WAREHOUSE Metodologia di progettazione di un Data Warehouse, scelta del processo, scelta della granularità, identificare e rendere conformi le dimensioni, scelta delle misure, memorizzare pre-calcoli nella tabella dei fatti, completare la tabella delle dimensioni, scelta della durata del database, tracciare le "slowly changing dimension", decidere le priorità sulle query e sulle modalità di query, integrazione dei Data Mart.
4	RICHIAMI DI DATA MINING Descrizione dei Concetti. Market Basket Analysis. Classificazione e Predizione. Alberi di decisione. K-Nearest Neighbor. Clustering.
2	MEMORIZZAZIONE STRUTTURATA Database non relazionali. Tipologie di database NoSQL. Analisi di vantaggi e

	svantaggi dei database non relazionali. Esempi di implementazione di database NoSQL.
4	ANALISI DI ALCUNI CONTESTI REALI DI BIG DATA Reti sociali. Reti biologiche. Next-Generation Sequencing (NGS).
24	PROGETTAZIONE E LABORATORIO Realizzazione di Database e applicazioni. Utilizzo di strumenti di ETL per la progettazione e realizzazione di Data Warehouse, OLAP, Data Mining.
TESTI CONSIGLIATI	M. Golfarelli, S. Rizzi, <i>“Data Warehouse – Teoria e pratica della progettazione”</i> , Seconda Edizione, McGraw Hill, 2005. J. Han, M. Kamber, <i>“Data Mining: Concepts and Techniques”</i> , Morgan Kaufmann Publishers, 2006. H. Garcia-Molina, J.D. Ullman, J. Widom, <i>“Database Systems – The complete book”</i> , Second Edition, Pearson Education, 2008.