



UNIVERSITÀ DEGLI STUDI DI PALERMO

DIPARTIMENTO	Matematica e Informatica
ANNO ACCADEMICO OFFERTA	2017/2018
ANNO ACCADEMICO EROGAZIONE	2017/2018
CORSO DILAUREA MAGISTRALE	INFORMATICA
INSEGNAMENTO	BIG DATA MANAGEMENT
TIPO DI ATTIVITA'	B
AMBITO	50341-Discipline Informatiche
CODICE INSEGNAMENTO	17393
SETTORI SCIENTIFICO-DISCIPLINARI	INF/01
DOCENTE RESPONSABILE	ROMBO SIMONA ESTER Professore Ordinario Univ. di PALERMO
ALTRI DOCENTI	
CFU	6
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	102
NUMERO DI ORE RISERVATE ALLA DIDATTICA ASSISTITA	48
PROPEDEUTICITA'	
MUTUAZIONI	
ANNO DI CORSO	1
PERIODO DELLE LEZIONI	2° semestre
MODALITA' DI FREQUENZA	Facoltativa
TIPO DI VALUTAZIONE	Voto in trentesimi
ORARIO DI RICEVIMENTO DEGLI STUDENTI	ROMBO SIMONA ESTER Lunedì 9:30 13:30 In presenza: Via Archirafi 34, Piano II, Stanza 220 - Telematico: via Microsoft Teams o altro canale - In entrambi i casi, e' consigliabile prenotarsi tramite email alla docente

DOCENTE: Prof.ssa SIMONA ESTER ROMBO

PREREQUISITI	Conoscenze di base sulla progettazione di basi di dati. Conoscenze avanzate di programmazione.
RISULTATI DI APPRENDIMENTO ATTESI	<p>Conoscenza e capacita' di comprensione</p> <p>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.</p> <p>Capacita' di applicare conoscenza e comprensione Capacita' 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.</p> <p>Autonomia di giudizio Capacita' di analizzare e valutare soluzioni per la gestione di grosse quantita' di dati. Capacita' di progettare sistemi di supporto alle decisioni mediante l'analisi delle specifiche tecniche fornite.</p> <p>Abilita' comunicative Capacita' di descrivere soluzioni di progettazione di sistemi informativi complessi e sistemi di supporto alle decisioni e analizzarne le prestazioni. Capacita' di cooperare per determinare delle soluzioni di progettazione appropriate in contesti applicativi caratterizzati dalla presenza di grosse quantita' di dati.</p> <p>Capacita' d'apprendimento Capacita' di aggiornamento attraverso la consultazione di testi avanzati e di pubblicazioni scientifiche correlate agli argomenti affrontati durante il corso. Capacita' di seguire, utilizzando le conoscenze acquisite, sia master di secondo livello, che corsi di approfondimento e seminari specialistici nell'ambito della Business Intelligence.</p>
VALUTAZIONE DELL'APPRENDIMENTO	<p>Prova progettuale, prova orale.</p> <p>Prova progettuale: consistera' nella realizzazione e implementazione di un progetto basato su tecnologie di Big Data. E' obbligatorio che il progetto sia realizzato da un team costituito da un minimo di due ad un massimo di cinque studenti, al fine di favorire lo sviluppo della capacita' di lavorare in team. Verranno effettuate due revisioni di progetto, durante le quali la docente verifichera' il contributo individuale di ciascun componente del team e assegnera' una valutazione alle diverse fasi di progettazione. Alla fine al progetto verra' attribuito un punteggio complessivo, pesato per ciascuno studente in funzione del contributo individuale. Contribuiranno ad una valutazione positiva l'originalita' della soluzione proposta, il rigore metodologico, la dimostrazione di avere acquisito appieno gli strumenti tecnici forniti dal corso. I requisiti per conseguire il punteggio minimo per accedere alla seconda parte dell'esame consistono nella capacita' di realizzare un progetto che, seppure in modo basilare, rispetti le specifiche assegnate e sia corretto.</p> <p>Prova orale: servira' per verificare le conoscenze acquisite durante il corso, l'autonomia nell'approfondire contenuti anche complessi e la capacita' individuale di trovare soluzioni a problemi proposti. La prova orale avra' inizio con la presentazione di un argomento a scelta tra quelli proposti dalla docente durante il corso, su cui lo studente avra' prodotto una breve tesina. Quindi, verra' verificata la capacita' critica e l'autonomia di giudizio dello studente attraverso un numero adeguato di domande. La valutazione della prova orale andra' a integrare il punteggio conseguito da ciascuno studente a seguito della valutazione del progetto. Si fara' la media aritmetica tra il voto conseguito per il progetto e quello conseguito per la prova orale. In particolare: 18-21: Conoscenza sufficiente di tutti gli argomenti del programma. 22-24: Conoscenza discreta di tutti gli argomenti del programma. 25-27: Conoscenza buona di tutti gli argomenti del programma. 28-30: Conoscenza ottima di tutti gli argomenti del programma.</p>
OBIETTIVI FORMATIVI	Obiettivo del corso di Big Data Management e' 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 quantita' di dati. Il corso iniziera' con una panoramica su Business Intelligence e si concentrera' su progettazione di Data Warehouse e OLAP, su cui verteranno anche parte delle attivita' di laboratorio. Il corso affrontera' anche la progettazione di database non relazionali e l'utilizzo di tecnologie quali Apache Hadoop e Spark. Infine, si mostreranno degli scenari applicativi reali quali ad esempio reti sociali, smart cities e dati biologici.
ORGANIZZAZIONE DELLA DIDATTICA	Lezioni frontali.
TESTI CONSIGLIATI	DATA WAREHOUSE: M. Golfarelli, S. Rizzi, "Data Warehouse – Teoria e pratica della progettazione", Seconda Edizione, McGraw Hill, 2005.

PROGRAMMA

ORE	Lezioni
2	INTRODUZIONE AI BIG DATA Le tre V. Esigenze. Contesti applicativi. Una panoramica sulle tecnologie.
6	BUSINESS INTELLIGENCE E DATA WAREHOUSING 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. Sistemi di supporto alle decisioni. Architetture per il Data Warehousing: architettura ad un livello, architettura a due livelli, architettura a tre livelli. I metadati. Qualita' 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 granularita, 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".
6	PROGETTAZIONE DI DATA WAREHOUSE IN PENTAHO Studio dei moduli: Kettle, Data Integration, OLAP.
4	ANALISI E VISUALIZZAZIONE Applicazioni e strumenti specifici per l'analisi di big data e la visualizzazione dei risultati di complesse query di tipo decisionale.
2	RICHIAMI DI DATA MINING Descrizione dei Concetti. Market Basket Analysis. Classificazione e Predizione. Alberi di decisione. K-Nearest Neighbor. Clustering.
6	MEMORIZZAZIONE STRUTTURATA Database non relazionali. Tipologie di database NoSQL. Analisi di vantaggi e svantaggi dei database non relazionali. Esempi di implementazione di database NoSQL ed esercizi.
6	TECNOLOGIE PER BIG DATA (I) Studio dell'algoritmo Map-Reduce. Studio di strumenti come Apache Hadoop e Spark. Applicazioni ed esercizi usando gli strumenti acquisiti.
4	TECNOLOGIE PER BIG DATA (II) Studio e applicazione di specifiche librerie di Spark.
6	ANALISI DI ALCUNI CONTESTI REALI DI BIG DATA Reti sociali. Reti biologiche. Next-Generation Sequencing (NGS). Smart Cities.