



UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Ingegneria
ACADEMIC YEAR	2015/2016
BACHELOR'S DEGREE (BSC)	MANAGEMENT AND COMPUTER ENGINEERING
SUBJECT	BIG DATA & ANALYTICS
TYPE OF EDUCATIONAL ACTIVITY	C
AMBIT	10655-Attività formative affini o integrative
CODE	18103
SCIENTIFIC SECTOR(S)	INF/01
HEAD PROFESSOR(S)	FERRARO PIERLUCA Ricercatore a tempo determinato Univ. di PALERMO
OTHER PROFESSOR(S)	
CREDITS	6
INDIVIDUAL STUDY (Hrs)	96
COURSE ACTIVITY (Hrs)	54
PROPAEDEUTICAL SUBJECTS	
MUTUALIZATION	
YEAR	3
TERM (SEMESTER)	2° semester
ATTENDANCE	Not mandatory
EVALUATION	Out of 30
TEACHER OFFICE HOURS	FERRARO PIERLUCA Tuesday 15:00 17:00 Edificio 6, terzo piano

DOCENTE: Prof. PIERLUCA FERRARO

PREREQUISITES	
LEARNING OUTCOMES	<p>Conoscenza e capacità di comprensione Lo studente al termine del Corso avrà conoscenza dei modelli di calcolo, delle architetture e delle infrastrutture necessarie all'elaborazione di grandi moli di dati. Avrà inoltre familiarità con le problematiche legate all'analisi di grandi moli di dati.</p> <p>Capacità di applicare conoscenza e comprensione Lo studente sarà in grado di analizzare scenari caratterizzati dalla presenza di grandi moli di data e sarà in grado di fornire opportune soluzioni progettuali per la realizzazione di sistemi in grado di gestire tali dati. Sarà in grado di progettare sistemi software efficienti per l'elaborazione di grandi moli di dati.</p> <p>Autonomia di giudizio Lo studente sarà in grado di valutare la bontà delle diverse soluzioni algoritmiche e architetture per la gestione di grandi moli di dati e sarà inoltre in grado di valutare le performance dei diversi approcci interpretando opportuni indicatori.</p> <p>Abilità comunicative Lo studente acquisirà la capacità di comunicare ed esprimere problematiche inerenti all'oggetto del corso. Sarà in grado di sostenere conversazioni su tematiche relative alla progettazione e gestione di sistemi per la gestione di grandi moli di dati.</p> <p>Capacità d'apprendimento Lo studente avrà appreso come analizzare differenti applicazioni che richiedono l'uso di Big Data e avrà acquisito gli strumenti necessari ad apprendere autonomamente i principi di funzionamento di nuovi tool per la gestione dei big data.</p>
ASSESSMENT METHODS	Prova Scritta, Prova orale.
EDUCATIONAL OBJECTIVES	L'obiettivo del modulo è di formare lo studente con riferimento alla conoscenza dei principi basilari dei sistemi di gestione di grandi moli di dati. In particolare sarà in grado di valutare diversi approcci architetture e algoritmici allo stato dell'arte e di individuare le soluzioni più adatte agli scenari sotto esame.
TEACHING METHODS	Lezioni, Esercitazioni.
SUGGESTED BIBLIOGRAPHY	Dispense fornite dal docente.

SYLLABUS

Hrs	Frontal teaching
2	Introduzione ai concetti di base dei Big Data: terminologia, aspetti principali ed esempi di applicazioni
4	Problematiche principali legate alla gestione dei Big Data: volume dei dati e occupazione di memoria, velocità dell'elaborazione e complessità computazionale, presenza di errori e accuratezza dei dati, comprensione dei dati.
5	Infrastrutture per la gestione dei Big Data: Architetture Distribuite e Parallele; Cloud Computing per i Big Data.
4	Memorizzazione dei Big Data: Memorizzazione strutturata; Database non relazionali; Tipologie di database NoSql;
2	Tecniche di pre-elaborazione dei Big Data: Tipi di errori; Gestione degli errori.
7	Tecniche di pre-elaborazione dei Big Data: Filtraggio, Trasformazione, Integrazione.
3	Riduzione del numero di dimensioni dei Big Data: Principal-Component Analysis, Singular-Value Decomposition.
3	Clustering dei Big Data: Partizionamento in Cluster; Fuzzy Clustering; Clustering Relazionale.
3	Algoritmi di classificazione dei Big Data: Criteri di classificazione; Classificatori Bayesiani; Support Vector Machines.
3	Algoritmi di Predizione basati sui Big Data: Macchine a stati finiti; Modelli probabilistici; Modelli Ricorrenti.
Hrs	Practice
6	Progettazione di sistemi per la gestione distribuita di big data
6	Progettazione di sistemi per l'elaborazione parallela di big data
6	Progettazione di sistemi per l'applicazione di algoritmi avanzati di analisi dei big data