



# UNIVERSITÀ DEGLI STUDI DI PALERMO

<b>DIPARTIMENTO</b>	Ingegneria
<b>ANNO ACCADEMICO OFFERTA</b>	2016/2017
<b>ANNO ACCADEMICO EROGAZIONE</b>	2018/2019
<b>CORSO DILAUREA</b>	INGEGNERIA GESTIONALE E INFORMATICA
<b>INSEGNAMENTO</b>	BIG DATA & ANALYTICS
<b>TIPO DI ATTIVITA'</b>	C
<b>AMBITO</b>	10655-Attività formative affini o integrative
<b>CODICE INSEGNAMENTO</b>	18103
<b>SETTORI SCIENTIFICO-DISCIPLINARI</b>	INF/01
<b>DOCENTE RESPONSABILE</b>	FERRARO PIERLUCA      Ricercatore a tempo determinato      Univ. di PALERMO
<b>ALTRI DOCENTI</b>	
<b>CFU</b>	6
<b>NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE</b>	96
<b>NUMERO DI ORE RISERVATE ALLA DIDATTICA ASSISTITA</b>	54
<b>PROPEDEUTICITA'</b>	
<b>MUTUAZIONI</b>	
<b>ANNO DI CORSO</b>	3
<b>PERIODO DELLE LEZIONI</b>	2° semestre
<b>MODALITA' DI FREQUENZA</b>	Facoltativa
<b>TIPO DI VALUTAZIONE</b>	Voto in trentesimi
<b>ORARIO DI RICEVIMENTO DEGLI STUDENTI</b>	<b>FERRARO PIERLUCA</b> Martedì    15:00    17:00    Edificio 6, terzo piano

**DOCENTE:** Prof. PIERLUCA FERRARO

<b>PREREQUISITI</b>	Calcolatori elettronici e Programmazione.
<b>RISULTATI DI APPRENDIMENTO ATTESI</b>	<p>Conoscenza e capacita' di comprensione Lo studente al termine del Corso avra' conoscenza dei modelli di calcolo, delle architetture e delle infrastrutture necessarie all'elaborazione di grandi moli di dati. Avra' inoltre familiarita' con le problematiche legate all'analisi di grandi moli di dati.</p> <p>Capacita' di applicare conoscenza e comprensione Lo studente sara' in grado di analizzare scenari caratterizzati dalla presenza di grandi moli di data e sara' in grado di fornire opportune soluzioni progettuali per la realizzazione di sistemi in grado di gestire tali dati. Sara' in grado di progettare sistemi software efficienti per l'elaborazione di grandi moli di dati.</p> <p>Autonomia di giudizio Lo studente sara' in grado di valutare la bonta' delle diverse soluzioni algoritmiche e architetture per la gestione di grandi moli di dati e sara' inoltre in grado di valutare le performance dei diversi approcci interpretando opportuni indicatori.</p> <p>Abilita' comunicative Lo studente acquisira' la capacita' di comunicare ed esprimere problematiche inerenti all'oggetto del corso. Sara' in grado di sostenere conversazioni su tematiche relative alla progettazione e gestione di sistemi per la gestione di grandi moli di dati.</p> <p>Capacita' d'apprendimento Lo studente avra' appreso come analizzare differenti applicazioni che richiedono l'uso di Big Data e avra' acquisito gli strumenti necessari ad apprendere autonomamente i principi di funzionamento di nuovi tool per la gestione dei big data.</p>
<b>VALUTAZIONE DELL'APPRENDIMENTO</b>	Le conoscenze e le competenze acquisite dallo studente saranno verificate attraverso una prova scritta e un colloquio orale. La prova scritta sara' costituita da almeno tre esercizi volti a verificare le conoscenze dello studente degli argomenti affrontati durante il corso, e di applicare le capacita' e le conoscenze acquisite. Durante il colloquio orale lo studente dovra' essere in grado di discutere le soluzioni proposte durante la prova scritta; inoltre saranno proposte domande di diverso e crescente livello di complessita' al fine di valutare il raggiungimento degli obiettivi formativi e le abilita' comunicative dello studente. Infine, allo scopo di valutare l'autonomia di giudizio, sara' richiesto di analizzare le caratteristiche di specifici scenari applicativi e di proporre le soluzioni piu' adeguate ai problemi individuati.
<b>OBIETTIVI FORMATIVI</b>	L'obiettivo del corso e' di formare lo studente con riferimento alla conoscenza dei principi basilari dei sistemi di gestione di grandi moli di dati. In particolare sara' in grado di valutare diversi approcci architetture e algoritmici allo stato dell'arte e di individuare le soluzioni piu' adatte agli scenari sotto esame.
<b>ORGANIZZAZIONE DELLA DIDATTICA</b>	Lezioni, Esercitazioni.
<b>TESTI CONSIGLIATI</b>	Dispense fornite dal docente.

### PROGRAMMA

ORE	Lezioni
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, velocita' dell'elaborazione e complessita' 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.
ORE	Esercitazioni
6	Progettazione di sistemi per la gestione distribuita di big data
6	Progettazione di sistemi per l'elaborazione parallela di big data

<b>ORE</b>	<b>Esercitazioni</b>
6	Progettazione di sistemi per l'applicazione di algoritmi avanzati di analisi dei big data