



ANNO ACCADEMICO: 2019/2020

INSEGNAMENTO/MODULO: BIG DATA

TIPOLOGIA DI ATTIVITÀ FORMATIVA: Caratterizzante

DOCENTE: CARLO SARTIANI

e-mail: carlo.sartiani@unibas.it

sito web:

telefono: 0971205862

cell. di servizio (facoltativo):

Lingua di insegnamento: Italiano

n. CFU: 6	n. ore: 48	Sede: Potenza Dipartimento/Scuola: Scuola di Ingegneria CdS: Ingegneria informatica e delle tecnologie dell'informazione	Semestre: II
-----------	------------	---	--------------

OBIETTIVI FORMATIVI E RISULTATI DI APPRENDIMENTO

- **Conoscenza e capacità di comprensione:** lo studente deve dimostrare di conoscere e saper comprendere le problematiche relative all'utilizzo dei sistemi di elaborazione di Big Data, nonché le problematiche relative all'analisi di dataset massivi.
- **Capacità di applicare conoscenza e comprensione:** lo studente deve dimostrare di essere in grado di progettare e/o realizzare applicazioni basate sui sistemi di elaborazione di Big Data. Lo studente, inoltre, deve dimostrare di essere in grado di applicare le proprie conoscenze per l'analisi dei dati, la valutazione della bontà di applicazioni esistenti, nonché per l'adattamento e reingegnerizzazione di quelle ritenute non più soddisfacenti.
- **Autonomia di giudizio:** lo studente deve essere in grado di valutare e elaborare in modo autonomo gli aspetti critici relativi ai sistemi di Big Data e all'analisi dei dati; in particolare, lo studente deve essere in grado di valutare dati e fabbisogni informativi per identificare la piattaforma di elaborazione più adeguata, integrare molteplici piattaforme e svolgere correttamente l'analisi dei dati.
- **Abilità comunicative:** lo studente deve essere in grado di comunicare con altri specialisti tecnico scientifici per interagire e cooperare con loro. Inoltre, deve essere in grado di comunicare con esperti del dominio per trasmettere le conoscenze acquisite attraverso l'analisi dei dati.
- **Capacità di apprendimento:** lo studente deve essere in grado di aggiornarsi continuamente, tramite la consultazione di testi e pubblicazioni scientifiche, in modo tale da adattarsi ai continui cambiamenti di un'area profondamente dinamica come quella dei Big Data. Lo studente, inoltre, deve essere in grado di frequentare corsi di data science presenti in master universitari o programmi di dottorato.

PREREQUISITI

È necessario avere acquisito e assimilato le seguenti conoscenze (fornite dal corso di Basi di Dati e dai corsi di programmazione):

- concetti elementari di programmazione procedurale e a oggetti;
- logica proposizionale e logica del primo ordine;
- basi di dati e SQL.



CONTENUTI DEL CORSO

Molte importanti decisioni vengono prese sulla base di analisi di grandi quantità di dati. Ciò accade in settori quali la scienza, la tecnologia, l'economia, la politica e la medicina.

Questo corso vuole fornire un'ampia introduzione alle problematiche connesse alla gestione e all'analisi di Big Data.

- Introduzione, logistica del corso, introduzione ai Big Data. (3h)
- Sistemi di gestione di Big Data: MapReduce, Hadoop, Spark (5-2h).
- Analisi e visualizzazione di dati attraverso spreadsheets (4-2h)
- Visualizzazione dati avanzata attraverso Tableau (2h).
- Discussione del progetto. Formazione dei gruppi. (2h).
- Introduzione a Python. Python per l'analisi e la visualizzazione dei dati. (6h-2h).
- Introduzione al Machine Learning. Python per il Machine Learning. (6h)
- Il linguaggio R: analisi dei dati, visualizzazione, machine learning. (2-2h)
- Algoritmi di Data Mining (2h).
- Data Mining con SQL e Python (2h)
- Analisi di grafi (2h).
- Analisi di dati non strutturati (2h)
- Big Data nel mondo reale. Introduzione al Deep Learning (2h)

METODI DIDATTICI

Il corso prevede 48 ore di didattica tra lezioni ed esercitazioni. In particolare sono previste 38 ore di lezione in aula e 10 ore di esercitazioni.

MODALITÀ DI VERIFICA DELL'APPRENDIMENTO

- Realizzazione di un progetto
- Partecipazione alle attività in aula
- Discussione del progetto

Sulla base delle modalità di esame previste per corsi simili in molte università italiane ed estere, gli studenti saranno suddivisi in gruppi (2-3 studenti) e dovranno svolgere un progetto su un argomento scelto ogni anno: tale progetto dovrà essere consegnato entro una deadline fissa. Ogni gruppo dovrà utilizzare sistemi e/o linguaggi diversi, garantendo il raggiungimento di obiettivi quali l'accuratezza dei risultati, la scalabilità e l'efficienza.

Gli studenti, inoltre, faranno parte di gruppi di supporto, che hanno l'obiettivo di assistere gli studenti stessi nello svolgimento del progetto. I componenti di ciascun gruppo, quindi, acquisiranno conoscenze estremamente approfondite su uno specifico argomento del corso.

Gli studenti che non hanno la possibilità di frequentare il corso, dovranno concordare un progetto con il docente e avranno 30 giorni di tempo per il suo completamento.

Il voto finale sarà determinato sulla base del lavoro svolto nei gruppi di supporto e sulla base della qualità del progetto realizzato.

TESTI DI RIFERIMENTO E DI APPROFONDIMENTO, MATERIALE DIDATTICO ON-LINE

Appunti forniti dal docente disponibili sul sito Moodle del corso.

METODI E MODALITÀ DI GESTIONE DEI RAPPORTI CON GLI STUDENTI

Orario di ricevimento: il martedì dalle ore 15,00 alle 17,00 presso lo studio del docente e il mercoledì dalle 15,00 alle 17,00 presso lo studio del docente.

Oltre all'orario di ricevimento settimanale, il docente è disponibile in ogni momento per un contatto con gli studenti attraverso la propria e-mail.

DATE DI ESAME PREVISTE¹

¹ Potrebbero subire variazioni: consultare la pagina web del docente o del Dipartimento/Scuola per eventuali aggiornamenti



Università degli Studi della Basilicata
Scuola di Ingegneria

-
- 01/07/2020
 - 14/07/2020
 - 30/09/2020
 - 16/12/2020

SEMINARI DI ESPERTI ESTERNI SI NO

ALTRE INFORMAZIONI

