STRUTTURA	Scuola Politecnica
ANNO ACCADEMICO	2015/2016
CORSO DI LAUREA MAGISTRALE	Scienze Statistiche (LM-82)
INSEGNAMENTO	Metodi Esplorativi e Data Mining
TIPO DI ATTIVITÀ	Caratterizzante
AMBITO DISCIPLINARE	Statistico
CODICE INSEGNAMENTO	14328
ARTICOLAZIONE IN MODULI	NO
SETTORI SCIENTIFICO DISCIPLINARI	SECS/S-01
DOCENTE RESPONSABILE	Antonella Plaia Professore Associato Università di Palermo
CFU	6
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	102
NUMERO DI ORE RISERVATE ALLE ATTIVITÀ DIDATTICHE ASSISTITE	42 (30 LF+12 Es/Lab)
PROPEDEUTICITÀ	Nessuna
ANNO DI CORSO	II
SEDE DI SVOLGIMENTO DELLE LEZIONI	Consultare il sitopolitecnica.unipa.it
ORGANIZZAZIONE DELLA DIDATTICA	Lezioni frontali, Esercitazioni
MODALITÀ DI FREQUENZA	Facoltativa
METODI DI VALUTAZIONE	Presentazione di una Tesina e Prova Orale
TIPO DI VALUTAZIONE	Voto in trentesimi
PERIODO DELLE LEZIONI	Primo semestre
CALENDARIO DELLE ATTIVITÀ DIDATTICHE	Consultare il sitopolitecnica.unipa.it
ORARIO DI RICEVIMENTO DEGLI STUDENTI	Consultare la pagine personale del docente

RISULTATI DI APPRENDIMENTO ATTESI

Conoscenza e capacità di comprensione

Conoscenza dei metodi e delle procedure statistiche per analisi esplorative di data-set di grandi dimensioni.

Capacità di applicare conoscenza e comprensione

Lo studente dovrà essere in grado di:

1. formulare correttamente un problema
2. scegliere soluzioni di analisi statistico-informatiche appropriate,
3. tradurre i risultati in decisioni operative.

Autonomia di giudizio

Lo studente dovrà essere in grado di:

1. tradurre in termini statistici una esigenza conoscitiva sorta in differenti campi applicativi
2. intervenire con attività di pulizia, riorganizzazione, analisi descrittiva e interpretazione, dei dati,
3. elaborare e comunicare coerentemente i risultati delle proprie analisi.

Abilità comunicative

Lo studente dovrà essere in grado di:

1. comunicare con chiarezza, sia agli specialisti che ai non specialisti della materia, i concetti, e le tecniche di analisi dei dati studiati,
2. giustificare le scelte degli strumenti utilizzati per l'analisi,
3. comunicare i risultati con linguaggio appropriato.

Capacità d'apprendimento

Lo studente avrà sviluppato le abilità di apprendimento necessarie per approfondire autonomamente lo studio delle tecniche statistiche più comunemente utilizzate nell'analisi di grandi moli di dati.

OBIETTIVI FORMATIVI DEL CORSO

Il corso illustra metodi statistici di apprendimento da dati empirici complessi. L'obiettivo principale è l'analisi di grandi database al fine di trovare pattern, associazioni, cambiamenti, anomalie e strutture di particolare interesse.

ORE FRONTALI	LEZIONI FRONTALI
2	Analisi esplorativa dei dati
2	Trasformazioni di variabili
6	Trattamento dei dati mancanti
2	Alberi decisionali
2	Tecniche di segmentazione
6	Alberi di classificazione
4	Alberi di regressione
	ESERCITAZIONI
2	Trasformazioni di variabili
8	Trattamento dei dati mancanti
8	Alberi di classificazione
6	Alberi di regressione
TESTI CONSIGLIATI	Dispense rese disponibili dal docente sul portale di Ateneo. Risorse on-line. P. Giudici, Data mining: metodi informatici, statistici ed applicazioni, II edizione, McGraw-Hill, Milano 2005 Breiman, L. Friedman, J. H. Olshen, R. A. Stone, C. J. (1984) Classification and regression trees, Chapman & Hall.