



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Informatica

1920-1-E0201Q046

Obiettivi

L'insegnamento si propone di spiegare i concetti base dell'Informatica e della Statistica, con una particolare attenzione alle applicazioni di carattere biotecnologico. Durante l'insegnamento verranno illustrate le relazioni fra le discipline informatiche e le discipline biologiche: la trattazione dei concetti base dell'Informatica sarà affiancata alla spiegazione dell'importanza che tali concetti rivestono per l'analisi di dati biologici (ad es. ricerca in banche dati biologiche, problemi di Bioinformatica, Biologia Computazionale e Biologia dei Sistemi).

Conoscenza e capacità di comprensione. Al termine dell'insegnamento lo studente saprà sviluppare la capacità di:

- "computational thinking", al fine di utilizzare in modo appropriato gli strumenti dell'Informatica (algoritmi, metodi computazionali, software) per la soluzione di un dato problema;

- analisi critica relativamente alla scelta dei metodi statistici più adeguati per l'analisi di dati in ambito biologico/clinico, e per l'interpretazione dei rispettivi risultati.

Capacità di applicare conoscenza e comprensione. Al termine dell'insegnamento lo studente sarà in grado di applicare le conoscenze acquisite per la soluzione di problemi in ambito computazionale e statistico:

- organizzare e trattare automaticamente i dati o risolvere semplici problemi tramite l'utilizzo di fogli di calcolo (esercitazioni in laboratorio);

- sviluppare semplici algoritmi, programmare in linguaggio Python e utilizzare l'ambiente di sviluppo R per analisi statistiche (esercitazioni in laboratorio).

Autonomia di giudizio. Al termine dell'insegnamento lo studente sarà in grado di elaborare quanto appreso, e saprà riconoscere le situazioni e i problemi in cui le metodologie informatiche e statistiche apprese possano essere utilizzate.

Abilità comunicative. Al termine dell'insegnamento lo studente sarà in grado di esprimersi in modo appropriato nella descrizione delle tematiche affrontate, con proprietà di linguaggio e sicurezza di esposizione.

Capacità di apprendimento. Al termine dell'insegnamento lo studente sarà in grado di consultare la letteratura sugli argomenti trattati, nonché analizzare, applicare, integrare e collegare le conoscenze acquisite con quanto verrà appreso in insegnamenti correlati all'applicazione delle scienze computazionali e statistiche per l'analisi dei dati.

Contenuti sintetici

Informatica

- 1) Nozioni di base sui calcolatori e sulla codifica dei dati
- 2) Pensare in modo algoritmico ed elementi di programmazione
- 3) Nozioni di Bioinformatica
- 4) Basi di dati

Statistica

- 1) Statistica descrittiva
- 2) Statistica inferenziale

Esercitazioni in laboratorio

Fogli di calcolo

Elementi di programmazione in Python e R

Programma esteso

Informatica

- 1) Nozioni di base sui calcolatori e sulla codifica dei dati
 - Cos'è un computer (architettura di von Neumann)
 - Differenza tra hardware e software (software applicativo e di sistema)
 - Cos'è un processore e quali istruzioni esegue
 - Cos'è la memoria e cosa contiene
 - Codifica dell'informazione (rappresentazione dei numeri e dei testi)
 - Rappresentazione dell'informazione multimediale
 - Visualizzazione e rappresentazione dei dati biologici
- 2) Pensare in modo algoritmico ed elementi di programmazione
 - Definizione di algoritmo
 - Dal problema all'algoritmo al programma
 - Programmi e linguaggi di programmazione
 - Programmazione strutturata e pseudo-codice
 - Strutture dati
 - Problemi "facili" e problemi "difficili"
 - Efficienza degli algoritmi (nozioni di complessità computazionale)
- 3) Nozioni di Bioinformatica
 - Problemi biologici e relative soluzioni informatiche
 - Le banche dati biologiche (banche dati primarie, secondarie, specializzate)
 - Ricerca in banche dati genomiche: algoritmi di allineamento e tecniche euristiche
 - Alcuni problemi "difficili" della Bioinformatica: protein folding, molecular docking
 - Nozioni di Biologia Computazionale e Biologia dei Sistemi
 - Dalla Biologia all'Informatica: metodi computazionali di ispirazione biologica per la soluzione di problemi "difficili"
- 4) Basi di dati
 - Differenze fra basi di dati e fogli di calcolo
 - Database management systems (DBMS)
 - Modello relazionale di una base di dati
 - Struttura delle banche dati biologiche

Statistica

- 1) Statistica descrittiva

Campioni e popolazioni, tipologie di dati e variabili
Disegno degli esperimenti (cieco, replicazione, strategie di campionamento)
Il concetto di frequenza (assoluta, relativa, cumulativa)
Rappresentazione grafica dei dati (istogrammi, diagrammi a torta, diagrammi a dispersione)
Misure di centralità (media, mediana, moda)
Misure di dispersione (range, deviazione standard, varianza)
Misure di posizione (quantili, percentili)
Analisi esplorativa dei dati (outlier, boxplot)
2) Statistica inferenziale
Nozioni di teoria della probabilità
Distribuzioni di probabilità (uniforme, binomiale, normale, Poisson)
Metodi di stima con un campione (intervalli di confidenza, stima di media e varianza con un campione, distribuzione t di Student, distribuzione chi-quadro)
Verifica di ipotesi con un campione
Correlazione e regressione

Esercitazioni in laboratorio
I fogli di calcolo per l'elaborazione dei dati
Strutture dati e ragionamento algoritmico
Elementi di programmazione in Python
Elementi di statistica descrittiva ed inferenziale con R

Prerequisiti

Prerequisiti: nessuno
Propedeuticità: nessuna

Modalità didattica

Lezioni frontali in aula (40 ore, 5 CFU) supportate da presentazioni PowerPoint, ed esercitazioni in laboratorio (30 ore, 3 CFU).

L'insegnamento è tenuto in lingua italiana.

Materiale didattico

Materiale didattico disponibile sulla piattaforma e-learning dell'insegnamento.

Libri di testo suggeriti:

- L. Snyder, A. Amoroso, Fluency - Conoscere e usare l'informatica, Pearson, 2015
- M.M. Triola, M.F. Triola, Fondamenti di statistica per le discipline biomediche, Pearson, 2017
- M.C. Whitlock, D. Schluter, Analisi statistica dei dati biologici, Zanichelli, 2010

Periodo di erogazione dell'insegnamento

Primo semestre

Modalità di verifica del profitto e valutazione

Esame scritto della durata di 2 ore, con 9 domande a risposta multipla e una domanda a risposta aperta, riguardanti gli argomenti trattati durante le lezioni frontali e le esercitazioni in laboratorio.

Orario di ricevimento

Ricevimento: su appuntamento, previa richiesta per mail al docente.
