

**CORSO DI LAUREA MAGISTRALE IN
BIOTECNOLOGIE PER L'APPROCCIO ONE HEALTH**

Informazioni Insegnamento

a.a. 2023/2024

Corso di BIG DATA

6 CFU, I anno, II semestre

SSD ING-INF/05

Informazioni Docente

Docente: Pietro Hiram Guzzi

indirizzo mail: hguzzi@unicz.it (5CFU)

Docente : Chiara Zucco, zucco@unicz.it (1CFU)

eventuali n.ri telefono per prendere appuntamenti, gli orari di ricevimento. Ricevimento su prenotazione da concordare a hguzzi@unicz.it sia in presenza che on line.

Descrizione del Corso

Il corso fornirà agli studenti una panoramica sulle principali tecniche di bioinformatica per l'analisi proteomica, genomica e trascrittomica. Allo studente saranno altresì trasmesse le competenze per l'utilizzo consapevole dei principali strumenti di analisi

Obiettivi del Corso e Risultati di apprendimento attesi

Lo studente a fine corso saprà impostare un esperimento in silico di analisi di dati proteomici, genomici, trascrittomici etc discernendo, in modalità problem solving, il giusto approccio a seconda del problema e del contesto. Come side-effects si auspica che lo studente possa realizzare delle pipeline di analisi autonome a partire dalle competenze acquisite.

Programma

- 1) Elementi di Informatica e statistica essenziale (8 ore): a) Elementi di architettura dei calcolatori, hw, sw di base e sw applicativo (anche reti e cloud) b) Algoritmi; potenza di calcolo e efficienza degli algoritmi c) Elementi di probabilità e statistica (Media, Mediana, Probabilità a priori e a posteriori, Teorema di Bayes, Verosimiglianza)
- 2) Organizzazione e gestione dei dati (8 ore):
 - a) Data base e DBMS: (Struttura dei DB, DB relazionali, progettazione e inserimento dati, metodologie di accesso e interrogazione)
 - b) Banche dati genomiche e proteomiche i) GenBank - ENA - DDBJ



((INSDC)): formato entry file, inserimento sequenze, ricerca sequenze
ii) Genome browsers: ENSEMBL, UCSC iii) UniprotKB (Swiss-Prot, TrEMBL)
iv) Esercitazioni: accesso e ricerca incrociata di sequenze e dati

3) L'analisi dei dati (20 ore): a) Confronto di sequenze (10 ore) i) Allineamento (locale o globale): definizioni; allineamento ottimo, punteggi di allineamento, Matrici di sostituzione ii) Algoritmi di allineamento: programmazione dinamica, algoritmi euristici (BLAST, FASTA) iii) Esercitazioni: ricerche per similarità in banche dati b) Filogenie e predizioni di patterns (10 ore)

i) Allineamenti multipli

ii) Costruzione di profili e predizione di patterns iii) Metodi di costruzione di alberi filogenetici iv) Esercitazioni: ricerche in banche dati di profili e patterns, allineamenti multipli

4) La bioinformatica strutturale (12 ore):

a) Predizione di strutture secondarie i) Metodi statistico-probabilistici (Chou e Fasman, GOR) ii) Metodi basati su intelligenza artificiale (reti neurali, HMM)
b) Predizione di strutture terziarie: i) Approcci template-based ii) Approcci template-free c) Interazioni proteina-ligandi: i) Docking ii) Computer-assisted drug design

Stima dell'impegno orario richiesto per lo studio individuale del programma

Il tempo stimato è di 150 ore, di cui 48 di attività frontali e 102 di studio individuale.

Metodi Insegnamento utilizzati

Il corso sarà organizzato in lezioni frontali, simulazione casi, problem solving, esercitazioni

Risorse per l'apprendimento

Libri di testo

Understanding Bioinformatics, 1 by Marketa Zvelebil

Pascarella Paiardini, Bioinformatica, Zanichelli

Ulteriori letture consigliate per approfondimento



Siti web delle società di bioinformatica.

Altro materiale didattico

(es. diapositive o dispense scaricabili dal sito)

Saranno fornite agli studenti le slide del corso sulla piattaforma di e learning.

Attività di supporto

Ricevimento studenti.

Modalità di frequenza

Le modalità sono indicate dal Regolamento didattico d'Ateneo.

Modalità di accertamento

Le modalità generali sono indicate nel regolamento didattico di Ateneo all'art.22 consultabile al link:

[http://www.unicz.it/pdf/regolamento didattico ateneo dr681.pdf](http://www.unicz.it/pdf/regolamento%20didattico%20ateneo%20dr681.pdf)

Durante il corso sarà svolto un esame in itinere in forma scritta

L'esame finale sarà svolto in forma scritta, con un test a risposta multipla composto da 15 domande, 2 punti a risposta corretta, l'esame sarà superato con un voto maggiore uguale a 18 (9 domande corrette).

