

Corso seminariale anno accademico 2022/23

SEMESTRE: secondo

CORSO DI LAUREA: triennale, magistrale

TITOLO CORSO: Metodi e modelli in Biofisica Computazionale (Laboratorio)

DOCENTI: Andrea BASCIU, Attilio Vittorio VARGIU

NUMERO CFU: 3

DESCRIZIONE/PROGRAMMA DEL CORSO:

Il laboratorio, rivolto agli studenti della laurea Triennale e Magistrale, vuole fornire una panoramica delle potenzialità dei metodi contemporanei di biofisica molecolare computazionale, inclusi alcuni algoritmi di intelligenza artificiale, applicati allo studio di problemi e processi di interesse biologico. Esso si articola in tre moduli:

Il primo consiste in una introduzione alla struttura e alla dinamica delle proteine, delle loro interazioni reciproche e con altre biomolecole. Durante il primo modulo, gli studenti familiarizzeranno con l'utilizzo di alcuni popolari programmi di visualizzazione di strutture e dinamiche di proteine e dei loro complessi con altre biomolecole.

Nel secondo modulo sarà fornita una panoramica su alcune tecniche di machine(deep) learning, supervisionato e non supervisionato, particolarmente utilizzate nell'analisi di dati biologici: cluster analysis e principal component analysis come metodi di apprendimento non supervisionato e support vector machines e random forest nel caso di apprendimento supervisionato. Per ogni argomento saranno presentati dei "case study" per mostrare delle possibili applicazioni dei concetti, e verranno proposte alcune semplici esercitazioni pratiche da svolgere in aula con gli studenti.

Il terzo modulo consiste in una introduzione al "docking molecolare", tecnica principe per lo studio del legame tra biomolecole e proteine, e di ampia applicazione in tutti i processi di sviluppo razionale di nuovi farmaci. Verranno presentati i più utilizzati algoritmi di docking, per ognuno dei quali verranno presentati diversi "case study" durante la lezione. Infine, verrà svolta un'esercitazione pratica sull'utilizzo di un software di docking.

Le conoscenze acquisite nei tre moduli verranno utilizzate per lo sviluppo di un progetto di machine learning basato su un set di dati proveniente da calcoli di docking molecolare.

Ogni modulo consiste di una parte teorica introduttiva e di una parte pratica più sostanziale, la quale si svolgerà in ambiente Linux, utilizzando i principali comandi della shell (bash, zsh), il programma di docking basato sul deep-learning "GNINA", i programmi di visualizzazione di strutture molecolari e proteiche (Marvin, VMD) ed il linguaggio di programmazione Python per la parte pratica sul machine learning, svolta sulla piattaforma Google Colab. Non sono necessari prerequisiti specifici né di programmazione/machine learning, né di biologia.

MODALITA' DI SVOLGIMENTO PREVISTA:

Sessioni pratiche presso il Laboratorio Computazionale, Dipartimento di Fisica – 1° piano Torre A, all'interno del quale saranno messe a disposizione dello studente delle workstation o laptop preconfigurati.

TESTI DI RIFERIMENTO:

Selezione di capitoli di libri; slides e tutorials redatti dai docenti

MODALITA' D'ESAME:

Prova pratica concernente l'applicazione dei metodi e dei programmi introdotti e utilizzati durante il corso.