

DENOMINAZIONE INSEGNAMENTO / ATTIVITÀ FORMATIVA

Strutture Dati per la Bioinformatica

Data Structures for Bioinformatics

INFORMAZIONI INSEGNAMENTO / ATTIVITÀ FORMATIVA

A.A.	2026/27	CdS	BIOINFORMATICA
Codice	8068440	Canale	unico
CFU	2	Lingua	Italiano

DOCENTE RESPONSABILE DELL'INSEGNAMENTO / ATTIVITÀ FORMATIVA

Andrea Guarracino

CODOCENTE

OBIETTIVI FORMATIVI E RISULTATI DI APPRENDIMENTO ATTESI

Italiano	OBIETTIVI FORMATIVI: L'insegnamento si propone di insegnare a lavorare con gli strumenti del mestiere in Python, a partire dallo sfruttare le peculiarità del linguaggio, all'organizzazione del codice fino alle librerie per la manipolazione dei dati e la visualizzazione dei risultati. CONOSCENZA E CAPACITÀ DI COMPrensIONE: Si richiede a studentesse e studenti la capacità di applicare le conoscenze acquisite all'analisi di dati. CAPACITÀ DI APPLICARE CONOSCENZA E COMPrensIONE: Studentesse e studenti devono dimostrare conoscenze e capacità di programmazione Python per l'analisi di dati. AUTONOMIA DI GIUDIZIO: Studentesse e studenti devono dimostrare capacità di valutare una soluzione di un problema bioinformatico nel contesto del tipo di analisi necessaria. ABILITÀ COMUNICATIVE: Studentesse e studenti devono dimostrare capacità comunicative inerenti ai risultati ottenuti dalle analisi condotte tramite la scrittura di script e l'utilizzo di grafici. CAPACITÀ DI APPRENDIMENTO: Studentesse e studenti devono dimostrare di aver sviluppato le capacità di apprendimento necessarie per approfondire in autonomia nuove librerie e strumenti di programmazione in Python, consultando documentazione tecnica e risorse online, così da affrontare in modo indipendente nuovi problemi di analisi dei dati in ambito bioinformatico e proseguire negli studi e nell'attività professionale.
----------	---

English	LEARNING OUTCOMES: <i>The course aims to teach how to work with day-to-day Python tools, starting from exploiting the peculiarities of the language, to the organization of the code up to the libraries for data manipulation and visualization of the results.</i> KNOWLEDGE AND UNDERSTANDING: <i>Students are required to have the ability to apply the knowledge acquired to data analysis.</i> APPLYING KNOWLEDGE AND UNDERSTANDING: <i>Students must demonstrate knowledge and skills of Python programming for data analysis.</i> MAKING JUDGEMENTS: <i>Students must demonstrate the ability to evaluate a solution to a bioinformatic problem in the context of the type of analysis needed.</i> COMMUNICATION SKILLS: <i>Students must demonstrate communication skills inherent to the results obtained from analyzes conducted through script writing and the use of graphs.</i> LEARNING SKILLS: <i>Students must demonstrate the learning skills needed to independently explore new Python libraries and programming tools, consulting technical documentation and online resources, so as to autonomously tackle new data analysis problems in bioinformatics and to continue their studies and professional activity.</i>
---------	---

PREREQUISITI

Italiano	Concetti base di programmazione (almeno un corso tra C e Ruby).
English	Basic programming concepts (at least C or Ruby).

PROGRAMMA E CRONOPROGRAMMA

Italiano	Fondamenti del linguaggio di programmazione Python; strutture dati utili alla bioinformatica; gestione di dati tramite librerie dedicate; analisi numerica tramite Numpy; libreria Pandas; graficare dati con Matplotlib; buone pratiche di programmazione, introduzione all'esplorazione di dati.
English	Python basics; data structures for bioinformatics; data management through dedicated packages; Numerical analysis with Numpy; Pandas package; Data plotting with Matplotlib; Introduction to Data Exploration.

TESTI ADOTTATI E BIBLIOGRAFIA DI RIFERIMENTO

Italiano	Il materiale didattico (notebook Jupyter) è disponibile nel repository https://github.com/AndreaGuarracino/DataStructuresForBioinformatics . Ogni argomento del programma è coperto dai notebook indicati di seguito:
----------	---

	- Fondamenti del linguaggio di programmazione Python (variabili, stringhe, slicing, liste, tuple, condizionali, cicli, funzioni, librerie): `Lesson1/Lesson1.ipynb`, `Lesson2/Lesson2.ipynb` - Strutture dati utili alla bioinformatica (dizionari, set, list comprehension): `Lesson2/Lesson2.ipynb` (sezione Dictionary I), `Lesson3/Lesson3.ipynb` - Gestione di dati tramite librerie dedicate (lettura/scrittura file, gestione di file e directory, ricerca di motivi): `Lesson2/Lesson2.ipynb` (sezione Files), `Lesson3/Lesson3.ipynb` (sezione File II), `Lesson5/Lesson5.ipynb` - Buone pratiche di programmazione (zip, enumerate, itertools, *args/**kwargs): `Lesson4/Lesson4.ipynb`, `Lesson8/Lesson8.ipynb` (sezione finale) - Analisi numerica tramite Numpy: `Lesson6/Lesson6.ipynb` (sezione Numpy) - Libreria Pandas: `Lesson6/Lesson6.ipynb` (sezione Pandas), `Lesson7/Lesson7.ipynb` - Graficare dati con Matplotlib (e Seaborn): `Lesson8/Lesson8.ipynb` - Introduzione all'esplorazione di dati: `Lesson7/Lesson7.ipynb` (sezione Aggregation and Grouping), `Lesson8/Lesson8.ipynb` Per ciascuna lezione è inoltre disponibile un notebook di esercizi (`ExercisesN.ipynb`). Bibliografia di riferimento: https://github.com/AndreaGuarracino/DataStructuresForBioinformatics
English	<i>The teaching material (Jupyter notebooks) is available in the repository https://github.com/AndreaGuarracino/DataStructuresForBioinformatics. Each topic of the program is covered by the notebooks listed below:</i> - <i>Python language fundamentals (variables, strings, slicing, lists, tuples, conditionals, loops, functions, libraries): `Lesson1/Lesson1.ipynb`, `Lesson2/Lesson2.ipynb`</i> - <i>Data structures for bioinformatics (dictionaries, sets, list comprehension): `Lesson2/Lesson2.ipynb` (Dictionary I), `Lesson3/Lesson3.ipynb`</i> - <i>Data management through dedicated packages (file I/O, file and directory handling, motif search): `Lesson2/Lesson2.ipynb` (Files), `Lesson3/Lesson3.ipynb` (File II), `Lesson5/Lesson5.ipynb`</i> - <i>Good programming practices (zip, enumerate, itertools, *args/**kwargs): `Lesson4/Lesson4.ipynb`, `Lesson8/Lesson8.ipynb` (final section)</i> - <i>Numerical analysis with Numpy: `Lesson6/Lesson6.ipynb` (Numpy section)</i> - <i>Pandas package: `Lesson6/Lesson6.ipynb` (Pandas section), `Lesson7/Lesson7.ipynb`</i> - <i>Data plotting with Matplotlib (and Seaborn): `Lesson8/Lesson8.ipynb`</i> - <i>Introduction to Data Exploration: `Lesson7/Lesson7.ipynb` (Aggregation and Grouping), `Lesson8/Lesson8.ipynb`</i> <i>An exercise notebook (`ExercisesN.ipynb`) is also provided for each lesson.</i>

Reference Bibliography:

<https://github.com/AndreaGuarracino/DataStructuresForBioinformatics>

DESCRIZIONE DELLE MODALITÀ DI SVOLGIMENTO E METODI DIDATTICI ADOTTATI

☐ In presenza ☒ A distanza

Italiano	Lezione Frontale
English	Lecture

MODALITÀ DI FREQUENZA

☐ frequenza obbligatoria ☒ frequenza facoltativa

DESCRIZIONE DELLE MODALITÀ DI FREQUENZA

Italiano	Settimanale
English	Weekly

MODALITÀ DI VALUTAZIONE

☐ Prova scritta ☒ Prova orale ☐ Prova di laboratorio ☒ Prova pratica
☐ Valutazione in itinere ☐ Valutazione di progetto ☐ Valutazione di tirocinio

COMPOSIZIONE DELLA COMMISSIONE VALUTATRICE

Andrea Guarracino

DESCRIZIONE DELLE MODALITÀ E DEI CRITERI DI VERIFICA DELL'APPRENDIMENTO

Italiano	Viene effettuato un esame orale che prevede lo svolgimento di una prova pratica sugli argomenti trattati durante il corso. Durante l'esame vengono posti problemi per saggiare la capacità di studentesse e studenti di risolverli e di esporre le soluzioni applicate. 18-21, studentesse e studenti hanno acquisito i concetti di base della disciplina, modo di esprimersi e linguaggio complessivamente corretti e appropriati. 22-25, studentesse e studenti hanno acquisito in maniera approfondita i concetti di base della disciplina, ed effettuano i collegamenti fra le varie materie. Presentano linearità nella strutturazione del discorso. Il linguaggio è appropriato e corretto.
----------	--

	26-29, studentesse e studenti possiedono un bagaglio di conoscenze completo e ben strutturato. Sono in grado di applicare e rielaborare in maniera autonoma, senza alcun errore, le conoscenze acquisite. Evidenziano ricchezza di riferimenti e capacità logico-analitiche con un linguaggio fluido, appropriato e vario. 30 e 30 e lode, studentesse e studenti possiedono un bagaglio di conoscenze completo e approfondito. Sanno applicare conoscenze a casi e problemi complessi ed estenderle a situazioni nuove. I riferimenti culturali sono ricchi e aggiornati. Si esprimono con brillante proprietà di linguaggio. Non idoneo: importanti carenze e/o inaccurately nella conoscenza e comprensione degli argomenti; limitate capacità di analisi e sintesi, frequenti generalizzazioni e limitate capacità critiche e di giudizio, gli argomenti sono esposti in modo non coerente e con linguaggio inappropriato. La prova orale ha una durata media di 20-30 minuti, comprensiva della prova pratica. Non sono previsti esoneri parziali.
English	<i>An oral exam is carried out which involves carrying out a practical test on the topics covered during the course. During the exam, problems are posed to test the students' ability to solve them and explain the applied solutions.</i> <i>18-21, students have acquired the basic concepts of the discipline, mode of expression and language that is generally correct and appropriate.</i> <i>22-25, students have acquired in depth the basic concepts of the discipline, and make connections between the various subjects, presenting linearity in the structuring of speech. The language is appropriate and correct.</i> <i>26-29, students have a complete and well-structured set of knowledge and are able to independently apply and re-elaborate the acquired knowledge without any errors, highlights a wealth of references and logical-analytical skills with a fluid, appropriate and varied language.</i> <i>30 and 30 cum laude, students have a complete and in-depth knowledge base, knows how to apply knowledge to complex cases and problems and extend it to new situations. The cultural references are rich and up-to-date and expresses himself with brilliant language properties.</i> <i>Failed: significant deficiencies and/or inaccuracies in knowledge and understanding of the topics; limited analysis and synthesis skills, frequent generalizations and limited critical and judgmental skills, the topics are presented in an inconsistent manner and with inappropriate language.</i>

	<i>The oral exam has an average duration of 20-30 minutes, including the practical test. No partial exemptions are foreseen.</i>
--	---