ANALYTICAL TOOLS IN APPLIED BIOLOGY (AB000004)

1. language

English

2. course contents

Coordinator: Prof. Alessandro ARCOVITO

Year Course: 1° year

Semester:1° semester

CFU/UFC: 10

Modules and lecturers:

- FUNDAMENTALS OF PROGRAMMING FOR BIOLOGICAL SCIENCES (AB000036) - 3 cfu - ssd ING-INF/05

Prof. Daniele TOTI

- PROTEIN MODELLING (AB000038) - 2 cfu - ssd BIO/10

Prof. Alessandro ARCOVITO

- BASIC OF IMAGING AND SPECTROSCOPY (AB000041) - 3 cfu - ssd NN

Prof. Gabriele FERRINI

- MACHINE LEARNING FOR BIOLOGICAL SCIENCES (AB000043) - 2 cfu - ssd ING-INF/05

Dr. Giada BIANCHETTI

3. bibliography

For the FUNDAMENTALS OF PROGRAMMING FOR BIOLOGICAL SCIENCES module, the material will be provided by the teacher and will consist of slides and manuals for further study.

For the PROTEIN MODELLING module, the material will be provided by the teacher and will consist of scientific articles.

For the BASIC OF IMAGING AND SPECTROSCOPY module, a reference book is Murphy, Douglas B., and Michael W. Davidson. *Fundamentals of light microscopy and electronic imaging*. John Wiley & Sons, 2012. The material will be provided by the teacher, consisting mainly in slides and scientific articles.

For the MACHINE LEARNING FOR BIOLOGICAL SCIENCES module, the material will be provided by the teacher. It will consist of slides and scientific papers.

4. learning objectives

This course aims to provide students a practical and theoretical knowledge of some of the analytical and quantitative tools applied in regenerative medicine to facilitate the understanding and analysis of biological phenomena. The course is focused on the key concepts of bioinformatics applied to the biological field, such as programming (different programming languages, designing and analyzing algorithms) and on the informatic tools used to predict the structure of proteins starting from their amino acid sequence up to obtain detailed information about their function. The course will also introduce the principles of different machine learning techniques and their applications in the biomedical field for the extrapolation of relevant information from large datasets (data mining and artificial intelligence). The course also includes a module focused on the various imaging techniques and mass spectrometry used in biomedical research to study morphology, molecular content, and processes in biological tissues, both in physiological and pathological conditions.

Knowledge and understanding

The course aims to provide the fundamental knowledge of analytical and quantitative tools to formally and computationally describe biological phenomena and provide models for the analysis and interpretation of experimental data. To know the fundamental tasks of bioinformatics such as basic programming, the use of software for the analysis of structures, biological networks, and machine learning methods applicable in regenerative medicine, and to understand the principles underpinning imaging and mass spectrometry.

Applying knowledge and understanding

Understand how programming, the use of software, and the techniques proposed can be employed to analyze and track data in the biomedical field and specifically in translational medicine. Being able to perform simulations in order to apply and implement the techniques acquired to real-world phenomena/datasets and understand how these can be exploited to investigate the therapeutic potential of stem cells.

Making judgements

Know the main analytical tools used in biomedical research and recognize the situations in which they can be used to expand the therapeutic potential of stem cells in regenerative medicine. Evaluate correct reasoning and applications and identify flaws in deductive and experimental processes.

Communication skills

Know how to communicate clearly and unambiguously, to correctly use technical language, to appropriately disseminate and expose both the principles underlying the methods addressed during the course and the results obtained from these techniques. To be able to generate and transmit questions, ideas, and solutions concerning the disciplines addressed, to specialist and non-specialist interlocutors.

Learning skills

The student must be able to demonstrate a good capacity for self-assessment, and to be continuously updated through scientific articles and online platforms (NCBI, ATCC, Human cell atlas, etc.), and scientific meetings.

5. PREREQUISITES

Basic knowledge of Physics, Chemistry, and Biochemistry is necessary, as they are integral parts of the undergraduate courses that grant access to this Master's degree program.

6. teaching methods

The teaching methods used in this course include frontal lectures with the aid of slides and video tutorials, as well as the opportunity to use computer labs for specific practical exercises to acquire principles of programming, software for the analysis of three-dimensional protein structures, and software for image analysis.

• Knowledge and understanding (Dublin 1): The teaching activities conducted with the help of slides and video tutorials will ensure an effective transfer of information.

• Applying knowledge and understanding (Dublin 2): The possibility of practical exercises will allow the assessment of applied knowledge acquisition.

• Making judgments (Dublin 3): Students will be encouraged to work on independent projects within the course, culminating in the completion of written papers that will be evaluated and contribute to the final grade.

• Communication skills (Dublin 4): Students will be invited to present case studies during lectures and showcase projects completed independently or in collaboration with other students to teachers and peers.

• Learning skills (Dublin 5): The opportunity to study autonomously and utilize modern software will ensure comprehensive and up-to-date training for students enrolled in this degree program.

7. other informationS

Professors receive students by appointment, also using the main remote interview platforms such as Zoom or Teams.

8. methods for verifying learning and for evaluation

The assessment of learning will take place according to the following procedure:

1. At the end of the module of FUNDAMENTALS OF PROGRAMMING FOR BIOLOGICAL SCIENCES, the student will take a partial test made up of open-ended programming problems to be solved in the Python language.

2. For the module of *MACHINE LEARNING FOR BIOLOGICAL SCIENCES*, the student will be required to complete a small project involving the analysis of biological data using models studied during the course, whose results will be presented to the classroom. The examination will include a presentation of the completed project and its results.

3. On the day of the exam, the student will take a written test consisting of a multiple-choice quiz of 30 questions (3 questions for each credit unit, with 5 possible answers and only 1 correct answer), to be completed in 30 minutes. To pass the written test, the student must answer at least 18 out of 30 questions correctly.

The final grade will be determined by the score of the quiz expressed on a scale of thirty, to which an evaluation of the previously undertaken partial test and of the project's presentation will be added. Each of them will contribute with a score of + or - 2 thirty. Therefore, the FINAL GRADE will be the GRADE OF THE WRITTEN TEST + or - 4 thirty. The minimum passing grade is 18/30, and the maximum grade is 30/30 with honors, which is achieved by obtaining at least 31/30 by combining the scores of the multiple-choice written test and the previously submitted papers. This examination method fully achieves the objectives of the Berlin descriptors regarding evaluation methods, as follows:

• Knowledge and understanding (Dublin 1): It is achieved by the opportunity to assess the student's ability to write a specific paper on course topics.

• Knowledge and understanding applied (Dublin 2): It is achieved through the ability to solve open-ended programming problems and answer a multiple-choice test on the entire course program.

• Judgement (Dublin 3): It is achieved through the ability to solve open-ended programming problems and answer a multiple-choice test on the entire course program.

• Communication skills (Dublin 4): It is achieved by the opportunity to assess the student's ability to write a specific paper on course topics.

• Learning skills (Dublin 5): It is achieved by combining different assessment methods into a single aggregated grade that measures this learning ability consistently.

9. program

< FUNDAMENTALS OF PROGRAMMING FOR BIOLOGICAL SCIENCES >

This course covers the following areas: (1) Basic Python syntax, assignments, variables, operators, and introduction to PyCharm IDE. (2) The definition, usage, and debugging of different types of functions, including Python’s built-in and custom libraries. (3) An exploration of data structures such as strings and collections, alongside the concept of 'objects' in Python. (4) Deeper understanding of function concepts, file, and memory management, alongside advanced exception handling. (5) Principles of object-oriented programming including classes, objects, abstraction, encapsulation, inheritance, polymorphism, and their implementation in Python. (6) Advanced object-oriented concepts, design patterns, and software architectures, alongside principles of recursion and date/time management. (7) Introduction to Numpy, Matplotlib, and Pandas libraries for scientific computation and data visualization. (8) Introduction to biomedical-focused libraries like RDKit and BioPython, including an overview of free biomedical research software.

< MACHINE LEARNING FOR BIOLOGICAL SCIENCES >

Introduction to Machine Learning: overview of machine learning and its applications in biological sciences; types of machine learning and basic concepts (feature space, training, testing, evaluation metrics). Biological data types and preprocessing: introduction to biological data types and data preprocessing techniques (handling missing values, outliers, normalization, standardization). Supervised learning algorithms (regression, classification) and practical applications to biological dataset, including images. Unsupervised learning approaches (clustering) and their application in the biomedical field. Fundamentals of deep learning: introduction to neural networks for drug discovery and image analysis.

< BASIC OF IMAGING AND SPECTROSCOPY >

Image formation and contrast generation in an optical microscope: light, lenses and geometrical optics, Abbe's theory, diffraction and the Airy disk, the point spread function, the spatial resolution. Phase contrast and darkfield microscopy. Polarization microscopy, Differential Interference Contrast (DIC) and modulation contrast microscopy. Fluorescence Microscopy (FM): physical basis of fluorescence, filters and the epi-illumination in FM, objectives and spatial resolution in FM, quenching, blinking, and photobleaching. Detection and image analysis: fundamentals of digital imaging, digital image processing. Specialized microscopies: confocal laser scanning microscopy, two-photon microscopy, superresolution imaging. Brief introduction to Transmission Electron Microscopy (TEM), Cryocrystallography, Mass Spectrometry (application: Mass Cytometry).

< PROTEIN MODELLING >

Theoretical models, introduction. Molecular models in 3D graphics: methodologies for representing molecules, atoms, and their properties; representations of molecular surfaces and volumes. Databases of three-dimensional and two-dimensional structures. Approaches to building molecular models: quantum approaches (overview, advantages, and limitations) and approaches based on classical physics. Force fields in describing molecular properties (all-atoms, united-atoms, and coarse-grained force fields). Methods for calculating the structure and properties of small molecules and proteins. Molecular docking: theoretical study of interactions between molecules; protein-ligand and protein-protein docking; most common algorithms in solving the docking problem. Molecular dynamics.