UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Informatica

2122-1-E0201Q046

Aims

The course will introduce the basic concepts of computer science and statistics, with a particular focus on biotechnology applications. The course will introduce the students to the relationships between computer science and biology by focusing on biological databases, and problems in bioinformatics, computational biology and systems biology.

1. Knowledge and understanding.

The student will gain knowledge about:
- “computational thinking”, in order to critically use concepts and tools of computer science (algorithms, computational methods, software) for the solution of a given problem;
- the choice of the proper sampling and statistical methods, and the interpretation of outcomes in the analysis of biological/clinical data.

2. Applying knowledge and understanding.

The student will be able to apply the knowledge listed in item 1 for the solution of computational and statistical problems:
- organizing and handling biological data in automatic ways (practical skills on the use of spreadsheets);
- development of basic algorithms using Python and R programming languages (hands-on sessions in pc room).

3. Making judgements.

The student will be able to process the acquired knowledge and choose the proper computational/statistical methods for different applications.

4. Communication skills.

Use of an appropriate scientific vocabulary and ability in oral/written reports

5. Learning skills.

Skills in literature reading and understanding, as well as in the elaboration, analysis and application of the acquired knowledge in other courses related to the application of computational and statistical methods for data analysis.

Contents

Informatics:
Introduction to computer science.
Algorithms, computational thinking, and basics of structured programming.
Notions of computational complexity.
DNA computing.
Fundamentals of bioinformatics, computational biology, systems biology.
Bio-inspired meta-heuristics.

Statistics:
Descriptive statistics.
Inferential statistics.

Laboratory:
Spreadsheets and biological databases.
Basic notions of programming in Python and R languages.

Detailed program

Informatics

1) Introduction to computer science. Principles of computer operations (von Neumann architecture, fetch-execute cycle). Data representation.
2) Computational thinking and basics of structured programming. Definition of algorithm. From problems to algorithms, from algorithms to programs. Program languages. Structured programming and pseudo-code. Notions of computational complexity. DNA computing: Adleman's experiment.
3) Fundamentals of bioinformatics, computational biology and systems biology. Biological databases. Sequence alignment: algorithms and heuristics. Protein folding, molecular docking. Computational approaches for complex biological systems.
4) From biology to computer science: bio-inspired computational methods, and their applications in bioinformatics.

Statistics

1) Descriptive statistics. Introduction to statistics (types of data, collecting sample data). Summarizing and graphing data (frequency distributions, histograms, boxplots). Statistics for describing, exploring and comparing data (measures of center, measures of variation, measures of relative standing).
2) Inferential statistics. Basic concepts of probability. Probability distributions (uniform, binomial, normal, t distribution). Sampling distributions and estimators. Hypothesis testing. Correlation and regression.

Laboratory

Spreadsheets. Searching biological databases. Basic notions of programming in Python and R languages.

Prerequisites

Background: none.
Prerequisites: none.

Teaching form

Classroom lectures (40 hours) and hands-on sessions (30 hours) supported by PowerPoint slides.
All lectures and hands-on sessions will be videorecorded and published on the Moodle platform.

Textbook and teaching resource

All the educational material will be available on Moodle platform.

Textbooks:

- M.G. Schneider, J.L. Gersting. Informatica. Algoritmi, architetture, linguaggi, applicazioni. Maggioli Editore - Apogeo Education, 2020
- P.J. Deitel, H.M. Deitel. Introduzione a Python. Per l'informatica e la data science. Pearson, 2020
- M. Helmer Citterich, F. Ferrè, G. Pavesi, C. Romualdi, G. Pesole. Fondamenti di Bioinformatica. Zanichelli, 2018
- 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

Semester

First semester.

Assessment method

Written exam (2 hours), consisting in 9 multiple-choice questions and 1 open question about the topics presented during the classroom lectures and the hands-on sessions.

No mid-term exams will be scheduled.

Office hours

On demand by e-email to daniela.besozzi@unimib.it