Bioinformatics (3MI09NAK34M)

I. GENERAL INFORMATION

Number of hours per semester: 2 lectures + 4 practices per week

Credits: 6

Semester: spring

Language: English

Prerequisites: basic mathematics and genetics

Course type: <u>mandatory</u>/optional

Department: Dpt. of Biometrics and Agricultural Informatics

Course leader: Dr. Ladányi, Márta PhD, associate professor, head of department

Course description: The aim of the course is to overview the basic methods of bioinformatics. At the beginning of the course, we provide a short statistical introduction then we learn some applications and databases. The main part of the course is focused on methods for sequence alignments, similarity searching, and phylogeny. Topics will be discussed with applications in a computer lab in a practical way with many examples. We use the statistical software RStudio.

Requirements: During the semester students have to
- write a report after every 6-hour tutorial session (lecture+practical) which should be uploaded to the e-learning system. (Deadline for the report upload: Sunday, 23:50)
- prepare a homework project (write an essay and present it): collect data, set scientific questions and goals based on the data, analyse and interpret the results, answer the asked questions with using the methods learned during the course (exact requirements are given during in the semester).

Assessment, grading: To get a signature, active participation is needed during the classes (by regularly uploading the reports to the e learning system in time) with maximum 2 absents. Grades are offered based on the quality of the weekly reports and homework project.

Course material, literature, recommended readings:
Special handouts (lecture notes, scripts, data files together with useful information are regularly uploaded to the e-learning system and are available during the course .

Jin Xiong: Essential Bioinformatics. Cambridge University Press. 2006

Discussed chapters:
1. Probability, descriptive statistics, distributions. Data structure, data cleaning.
2. Estimations, hypothesis testing, one- and two-sample parametric tests
3. Crosstabulation, Normality tests
4. ANOVA
5. Linear and nonlinear regression
6. Clusteranalysis
7. Bioinformatical databases
8. Pairwise local and global sequence alignment
9. Similarity searching
10. Special methods: UPGMA, NJ,
11. Phylogeny (parsimony)
12. Multiple sequence alignment
13. Hidden Markov Models
14. Special methods: Maximum likelihood, Bayesian, MCMC and Jacknife
15. Motif and Domain searching

Learning outcomes: After having completed the course, students will be able to manage and evaluate data statistically choosing the appropriate method correctly, moreover, to report the results in a suitable manner. They learn write and use basic scripts in RStudio. They learn sequence alignment, similarity searching and phylogeny methods using free websites combined with RStudio. They can apply their skills in reading scientific papers and they learn how to report, present and reason their own findings and conclusions professionally.

Attendance policy: Missing more than two tutorial sessions will result in loss of credit for the module. Please note that the two absences are provided for sickness, so save your absences for situations you really need them.

Programme: The course is for MSc in Biotechnology students.