Course Specifications

Valid as from the academic year 2016-2017

Bioinformatics (O000105)

Course size *(nominal values; actual values may depend on programme)*

Credits 5.0	**Study time 150 h**	**Contact hrs**	45.0 h

Course offerings and teaching methods in academic year 2016-2017

A (semester 1)	seminar: practical PC room classes	30.0 h
	lecture	15.0 h

Lecturers in academic year 2016-2017

De Neve, Wesley	TW06	lecturer-in-charge
Van Messem, Arnout	WE02	co-lecturer

Offered in the following programmes in 2016-2017	crdts	offering
Bachelor of Science in Molecular Biotechnology	5	A

Teaching languages

English

Keywords

Algorithms, Bioinformatics, Command line, Computing, Data analysis, Linux, Problem solving, Programming, Python, Scripting

Position of the course

Bioinformatics, sometimes called The New Biology, is an interdisciplinary field of study that develops tools and algorithms for analyzing biological data, combining techniques from the domain of computer science, statistics, mathematics, and engineering. It has evolved into a basic scientific discipline that is essential not only for fundamental biological research but also for medical diagnosis, cancer research, drug development, and agriculture.
Recent technological advances have led to the generation of huge amounts of biotech data at an ever-increasing pace. This data avalanche has created an urgent need in the field of biotech for scientists and engineers with strong computational skills. As a result, this course aims at introducing students to some of the fundamental tools and algorithms that are at the core of computational pipelines for automated processing of biotech data.

Contents

This course deals with selected computational topics, focusing on automated processing of biotech data. These selected computational topics may include, but are not limited to:

- principles of Unix-based operating systems;
- interactive command line usage;
- shell scripting and regular expressions;
- databases and SQL;
- cloud computing and Web development;
- an introduction to Big Data; and
- machine learning algorithms.

In addition, the automated processing of biotech data may include, but is not limited to, genomic data compression, splice site prediction, prediction of drug-target interaction, and gene expression analysis.

Initial competences

Knowledge of a high-level programming language.
Knowledge of algorithmic design and analysis.

Final competences

The student is able (1) to make use of Unix-based operating systems, (2) to automate the processing of biotech data through interactive usage of the command line and shell scripting, (3) to communicate with biotech databases through the use of SQL, and (4) to apply machine learning techniques to biotech data. Furthermore, the student is able (1) to setup and operate computational pipelines for automated processing of biotech data, deploying bioinformatics tools and algorithms based on either explicit programming or machine learning, and (2) to subsequently interpret the obtained results.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Lecture, seminar: practical PC room classes

Learning materials and price

Handbook: Mark G. Sobell (2012). A Practical Guide to Linux: Commands, Editors, and Shell Programming. Second edition. Prentice Hall, ISBN-13: 978-0133085044.
Slides shown during the lectures will be made available on Minerva, together with additional learning materials (e.g., background information and links to relevant websites). Digital tools like Eclipse and Jupyter Notebooks for writing source code. Students are required to have a personal laptop for use in this course.

References

Mark G. Sobell (2012). A Practical Guide to Linux: Commands, Editors, and Shell Programming. Second Edition. Prentice Hall, ISBN-13: 978-0133085044.
William Punch and Richard Enbody (2012). The Practice of Computing using Python. Second Edition. Addison Wesley, ISBN-13: 978-0136110675.
Neil C. Jones and Pavel A. Pevzner (2004). An Introduction to Bioinformatics Algorithms. The MIT Press, ISBN-13: 978-0262101066.
Phillip Compeau and Pavel Pevzner (2014). Bioinformatics Algorithms: An Active Learning Approach. Active Learning Publishers, ISBN-13: 978-0990374602.
Christopher M. Bishop (2007). Pattern Recognition and Machine Learning. Springer, ISBN-13: 978-0387310732.
Steven Haddock and Casey Dunn (2010). Practical Computing for Biologists. First Edition. Sinauer Associates, Inc, ISBN-13: 978-0878933914.
Ashley Shade, Tracy K. Teal (2015). Computing Workflows for Biologists: A Roadmap. PLOS Biology.
Pavel A. Pevzner (2004). Educating Biologists in the 21st Century: Bioinformatics Scientists versus Bioinformatics Technicians. Bioinformatics, Vol.20, No.14, pages 2159–2161.
Alejandra J. Magana, Manaz Taleyarkhan, Daniela Rivera Alvarado, Michael Kane, John Springer, and Kari Clase (2014). A Survey of Scholarly Literature Describing the Field of Bioinformatics Education and Bioinformatics Educational Research. CBE—Life Sciences Education, Vol. 13, pages 607–623.

Course content-related study coaching

Evaluation methods

end-of-term evaluation and continuous assessment

Examination methods in case of periodic evaluation during the first examination period

Written examination with open questions, skills test

Examination methods in case of periodic evaluation during the second examination period

Examination methods in case of permanent evaluation

Assignment

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible in modified form

Extra information on the examination methods

The non-periodic evaluation cannot be retaken during the second examination period. Therefore, the final score for the second examination period is computed twice. The first computation takes into account both the score of the non-periodic evaluation (that is, the score obtained during the first examination period, on a maximum of 8) and the

score of the second examination period (on a maximum of 12). The second computation only takes into account the score of the second examination period (on a maximum of 20). The final score for the second examination period is then equal to the maximum of the above two computations.

Calculation of the examination mark

Periodic evaluation (60%) + non-periodic evaluation (20% hands-on sessions + 20% project). To qualify for passing this course, both the score of the periodic and the non-periodic evaluation should be at least equal to 8/20. If that is not the case, the total course score will be subject to an upper limit of 7/20.