COURSE DATA

DATA SUBJECT

Code: 42591
Name: Programming and advanced computational techniques in bioinformatics
Cycle: Master's Degree
ECTS Credits: 3
Academic year: 2026-27

STUDY (S)

Degree	Center	Acad. year	Period
2116 - Master's Degree in Bioinformatics	Escola Tècnica Superior d'Enginyeria	2	First quarter

SUBJECT-MATTER

Degree	Subject-matter	Character
2116 - Master's Degree in Bioinformatics	Programming and advanced computational techniques in bioinformatics	COMPULSORY

COORDINATION

FERRIS CASTELL RICARDO

DIAZ VILLANUEVA WLADIMIRO

SUMMARY

This course explores the potential parallel programming can bring us to solve large problems bioinformatics.
Additionally, applications and libraries using the most widely used bioinformatics programming languages seen in the master.
Knowing of the main tools available in the scientific community for data storage and processing bioinformatics. You will see that is Hadoop, NoSQL, Big Data and Cloud Computing.

PREVIOUS KNOWLEDGE

RELATIONSHIP TO OTHER SUBJECTS OF THE SAME DEGREE

There are no specified enrollment restrictions with other subjects of the curriculum.

OTHER REQUIREMENTS

None.

COMPETENCES / LEARNING OUTCOMES

2116 - Master's Degree in Bioinformatics

Be able to access the information required (databases, scientific articles, etc.) and to interpret and use it sensibly.

Be able to access to information tools in other areas of knowledge and use them properly.

Comprender en qué tipo de aplicaciones la programación paralela y los grandes sistemas de computación son requeridos para la resolución de problemas bioinformáticos y analizar sus prestaciones.

Conocer y emplear las principales aplicaciones bioinformáticas y las librerías existentes para los lenguajes de programación vistos en el Máster.

Desarrollar la iniciativa personal y ser capaces de realizar una toma rápida y eficaz de decisiones en su labor profesional y/o investigadora.

Students should apply acquired knowledge to solve problems in unfamiliar contexts within their field of study, including multidisciplinary scenarios.

Students should be able to integrate knowledge and address the complexity of making informed judgments based on incomplete or limited information, including reflections on the social and ethical responsibilities associated with the application of their knowledge and judgments.

Students should communicate conclusions and underlying knowledge clearly and unambiguously to both specialized and non-specialized audiences.

Students should demonstrate self-directed learning skills for continued academic growth.

Students should possess and understand foundational knowledge that enables original thinking and research in the field.

To be able to assess the need to complete the scientific, historical, language, informatics, literature, ethics, social and human background in general, attending conferences, courses or doing complementary activities, self-assessing the contribution of these activities towards a comprehensive development.

Trabajar en equipo con eficiencia en su labor profesional y/o investigadora y con personas de diferente procedencia.

DESCRIPTION OF CONTENTS

1. BioPerl, Bioconductor, and Introduction to Pipelines.
The basic concepts of various useful tools for Bioinformatics will be presented, as well as an introduction to pipelines.

2. Python Bioinformatics Libraries
The basic concepts of various useful tools for Bioinformatics will be presented, as well as an introduction to pipelines.

3. Introduction to Parallel Programming.
The basic concepts of HPC will be presented, such as OpenMP, Cuda, and SSE. The Python language will be used for parallel programming.

4. Hadoop
Apache Hadoop is a software framework that supports distributed applications under a free license. 1 It allows applications to work with thousands of nodes and petabytes of data. Hadoop was inspired by the Google documents for MapReduce and Google File System (GFS).

5. Big Data.
In the information and communications technology sector, Big Data refers to systems that manipulate large data sets. The most common difficulties in these cases focus on data capture, storage, search, sharing, analysis, and visualization. Bioinformatics is one of the disciplines considered Big Data.

6. Cloud computing.
Cloud computing, also known as cloud services, cloud computing, or cloud computing, is a paradigm that allows computing services to be offered over the Internet.

7. NoSQL.
A NoSQL database provides a data storage and retrieval mechanism that uses more flexible consistency models than traditional relational databases.

WORKLOAD

PRESENCIAL ACTIVITIES

Activity	Hours
Theory	10,00
Laboratory	5,00
Total hours	**15,00**

NON PRESENCIAL ACTIVITIES

Activity	Hours
Attendance at other activities	4,00
Individual or group project	2,00
Independent study and work	20,00
Preparation of lessons	15,00
Preparation for assessment activities	10,00
Resolution of case studies	10,00
Total hours	**61,00**

TEACHING METHODOLOGY

MD1 - Task training of the teaching-learning environment interaction in the classroom through expository sessions. Previous assignments include preparation (information search, reading texts supplied by teachers), teaching sessions themselves and the later work of deepening.

MD2 - Learning through problem solving and case studies, through which it is acquiring skills on different aspects of materials and subjects.

MD3 - Actividades de laboratorio driving range. INCLUDE preparations realization of the con las seguimiento driving range and I support the teacher, freestanding trabajo y elaborations on-line reports las driving range.

MD4 - Cross-disciplinary skills. Include attendance at courses, conferences or round tables organized by the CEC of the Master and / or conduct of a bibliographic work on issues that contribute to the integral. It produces a report of activities.

EVALUATION

Continuous assessment of student interaction in the classroom or lab or online activities conducted in groups (10%).

Evaluation reports or reports submitted regarding training activities and case studies problems of cross-cutting activities or others that arise in individual form (20%).

Evaluation reports or reports submitted concerning laboratory practice (30%).

Evaluation of objective individual tests (one or more) (40%).

You need to get at least 5.0 on the evaluation of objective individual tests to make the average of the notes and have delivered at least 90% of the work.

On second call the weights of the different sections and requirements, being able to improve or deliver all jobs except those conducted in groups will be retained.

REFERENCES

- Parallel Processing via MPI & OpenMP, M. Firuziaan, O. Nommensen. Linux Enterprise, 10/2002
- Big Data: A Revolution That Will Transform How We Live, Work, and Think. Viktor Mayer-Schonberger, Kenneth Cukier.
- INTRODUCCION A LA PROGRAMACION PARALELA. FRANCISCO ALMEIDA , EDICIONES PARANINFO, S.A., 2008.

- Parallel and High Performance Programming with Python. Fabio Nelli (2023)