

**COURSE DATA****Data Subject**

<b>Code</b>	42585
<b>Name</b>	Algorithms in bioinformatics
<b>Cycle</b>	Master's degree
<b>ECTS Credits</b>	3.0
<b>Academic year</b>	2019 - 2020

**Study (s)**

<b>Degree</b>	<b>Center</b>	<b>Acad. year</b>	<b>Period</b>
2116 - Master's Degree in Bioinformatics	School of Engineering	1	Second term

**Subject-matter**

<b>Degree</b>	<b>Subject-matter</b>	<b>Character</b>
2116 - Master's Degree in Bioinformatics	2 - Algorithms in bioinformatics	Obligatory

**Coordination**

<b>Name</b>	<b>Department</b>
ARNAU LLOMBART, VICENTE	240 - Computer Science

**SUMMARY**

We will study the algorithms used in Bioinformatics, focusing mainly on sequence analysis and searching for patterns. It will analyze the management of biological databases and algorithms used to extract useful information from them. It will consider how to represent biological knowledge in an ontology. Algorithms will be introduced for use in systems biology, such as graphs and Bayesian networks. We will study notions of optimization methods as genetic algorithms.

**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 (RD 1393/2007) // LEARNING OUTCOMES (RD 822/2021)

### 2116 - Master's Degree in Bioinformatics

- 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.
- Be able to access the information required (databases, scientific articles, etc.) and to interpret and use it sensibly.
- Students should possess and understand foundational knowledge that enables original thinking and research in the field.
- Be able to access to information tools in other areas of knowledge and use them properly.
- 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.
- Desarrollar la iniciativa personal y ser capaces de realizar una toma rápida y eficaz de decisiones en su labor profesional y/o investigadora.
- Trabajar en equipo con eficiencia en su labor profesional y/o investigadora y con personas de diferente procedencia.
- Conocer, comprender y aplicar las bases algorítmicas de los problemas más comunes en bioinformática.

## LEARNING OUTCOMES (RD 1393/2007) // NO CONTENT (RD 822/2021)

Understand the basic algorithmic most common problems in bioinformatics (alignment and pattern matching).

Knowing systems biology algorithms, mainly involving the management of networks and complex data structures.



Understand and apply the basis of various optimization algorithms.

## DESCRIPTION OF CONTENTS

### 1. Introduction to algorithmic

Will present the basic content of the course. They present the main algorithms to study in this subject and its applications in Bioinformatics. It will analyze the efficiency of the various algorithms and algorithm design strategies.

Finally, we discuss the most commonly used data structures and algorithms we will see some examples.

### 2. Finding patterns and data mining.

Finding patterns and data mining in bioinformatics. The large volume of data that has the necessary bioinformatics search using specific methods.

### 3. Graphs

We will use graph theory to represent bioinformatics data. We will see example of use.

### 4. Methods for similarity search and alignment.

A sequence alignment in bioinformatics is a way to represent and compare two or more sequences to highlight their areas of similarity, which could indicate functional relationships or evolutionary relationships among genes or proteins consulted.

We'll see how it performs this technique.

### 5. Hidden Markov Models

Use of Markov Chains in bioinformatics. Example of use.

### 6. Ontologies.

In computer science and information science, an ontology formally represents knowledge as a set of concepts within a domain, and the relationships among those concepts. We will used ontology in Bioinformatics.



## 7. Bayesian Networks

A Bayesian network is a multivariate probability model that relates a set of random variables using a directed graph explicitly causal influence. Bayesian networks are an extremely useful tool in estimating probabilities to new evidence and use in bioinformatics.

## 8. Genetic algorithms

In the computer science field of artificial intelligence, a genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.

We will use GA in bioinformatics.

## WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	15,00	100
Development of group work	5,00	0
Development of individual work	10,00	0
Study and independent work	15,00	0
Readings supplementary material	10,00	0
Preparation of evaluation activities	10,00	0
Preparing lectures	5,00	0
Preparation of practical classes and problem	5,00	0
<b>TOTAL</b>	<b>75,00</b>	

## 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.

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

In the two calls:

SE1 Continuous assessment: minimum 5 and maximum 15.

SE2 Activities: minimum 50 and maximum 90.

SE4 Exams: minimum 20 and maximum 50.

## REFERENCES

### Basic

- Referencia b1: Algoritmos de Bioinformática: técnicas y aplicaciones. Ion Mandoiu , Alexander Zelikovsky. ISBN: 978-0-470-09773-1. Willey series in Bioinformatics. 2008.
- An introduction to Algorithms. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein. The MIT Press, Cambridge, Massachusetts London, England. 2009.
- An introduction to Bioinformatics Algorithms. N.C. Jones, D.A. Pevzner. The MIT Press, Cambridge, 2004.

### Additional

- Referencia c1: STRUCTURAL BIOINFORMATICS: An Algorithmic Approach, by Forbes Burkowski, University of Waterloo, Canada. CRC Press, a Chapman & hall book.

## ADDENDUM COVID-19

**This addendum will only be activated if the health situation requires so and with the prior agreement of the Governing Council**

### 1. Contenidos

(NO HAY CAMBIOS)

### 2. Volumen de trabajo y planificación temporal de la docencia

(NO HAY CAMBIOS)



### **3. Metodología docente**

Las clases han sido presenciales hasta el confinamiento. Posteriormente se han dado mediante videoconferencia síncrona BBC.

Se han subido presentaciones con audio incorporado. Se han incluido tutorías por videoconferencia.

### **4. Evaluación**

Antes la evaluación era 50% trabajos entregables, 20% presentación y 30% un examen.

Ahora, ante la imposibilidad de realizar un examen, ha pasado a ser 70% trabajos entregables, 30% presentación.

### **5. Bibliografía**

(NO HAY CAMBIOS)