



## IDENTIFYING DATA

### Knowledge Discovery

Subject	Knowledge Discovery			
Code	V02M123V01113			
Study programme	(*)Máster Universitario en Ciencias Biológicas: Biología Molecular, Computacional e Ambiental e Bio-Industrias			
Descriptors	ECTS Credits	Type	Year	Quadmester
	6	Optional	1st	1st
Language	English			
Department				
Coordinator	Fernández Riverola, Florentino			
Lecturers	Fernández Riverola, Florentino García Lourenco, Analia María Olivieri Cecchi, David Nicholas			
E-mail	riverola@uvigo.es			
Web	<a href="http://webs.uvigo.es/biologicalsciences/">http://webs.uvigo.es/biologicalsciences/</a>			
General description	The aims of the subject are: (1) extract, summarize and represent biological information, (2) apply machine learning techniques, and (3) establish and show relations between biological observations.			

## Competencies

Type A	Code	Competences Specific
	A1	(*)To know the scientific method and the correct use of the scientific terminology as well as to acknowledge the contribution that scientific research provides to the overall knowledge and professional practice.
	A2	(*)Ability to describe and to analyse biological diversity, the mechanisms determining the interactions with the biotic and abiotic environment and being able to select those which might have technical applications.
	A3	(*)Ability to manage and/or to develop basic tools for validating and analysing data by means of statistics and bioinformatics.
	A4	(*)To know the ethical and legal aspects governing the collection and the handling of biological samples, organisms and habitats.
	A5	(*)Ability to design, evaluate and implement models of biological structures, systems and processes.
	A6	(*)To learn the sampling techniques and the instrumental methodologies, in the field and laboratory, for their application in the Biological Sciences
	A7	(*)To have an integrated view of the R&D processes and their possible transfer to the industrial sector. Planning and supervising facilities together with managing their human and economic resources.
	A8	(*)Ability to classify, evaluate, conserve, restore and manage natural and productive systems. Developing and implementing land management and sustainability plans.
	A9	(*)To understand and know how to apply quality control systems and safety protocols in any biological laboratory of the public or private sector.
	A10	(*)To acquire the professional ability to teach and spread Biology and to offer expertise advice for elaborating scientific, technical and socioeconomic biology reports. Address environmental consulting.

A11 (\*)To perform an individual Master Project (critical and in-depth study) under the supervision of a tutor in a research or working environment demonstrating that skills have been acquired.

#### Type B Code Competences Transversal

B1	(*)Dissemination of results and conclusions of the biological studies, in oral and written English, through complex presentations that address ideas related with R&D in Biology.
B2	(*)Managing computational, laboratory, field and industrial techniques in order to obtain, process and apply the acquired information.
B3	(*)Disseminating and broadcasting ideas in contexts both academic and non-specialised.
B4	(*)Reflecting on social and ethical responsibilities.

#### Learning aims

Subject competences	Typology	Competences
Ability to manage and/or to develop basic tools for validating and analysing data by means of statistics and bioinformatics.	know Know How	A3
bility to design, evaluate and implement models of biological structures, systems and processes.	know Know How	A5
To learn the sampling techniques and the instrumental methodologies, in the field and laboratory, for their application in the Biological Sciences.	know Know How	A6
Dissemination of results and conclusions of the biological studies, in oral and written English, through complex presentations that address ideas related with R&D in Biology.	know Know How	B1
Managing computational, laboratory, field and industrial techniques in order to obtain, process and apply the acquired information.	know Know How	B2
Disseminating and broadcasting ideas in contexts both academic and non-specialised.	know Know How	B3

#### Contents

Topic	
Biological databases	Common databases Data access Data interoperability
Data Mining/Machine Learning	Basic concepts Introduction to main approaches
Applications of Machine Learning to Sequence and Genome Analysis	Use of graphical probabilistic models for inference Associative Rule Extraction

#### Planning

	Personalized attention	Assessment	Ordinary class hours A	Face-to-face hours outside the classroom Guided academic environment B	Student's work factor C	Outside the classroom hours D	Total hours (A+B+D) E
Master Session	<input type="checkbox"/>	<input checked="" type="checkbox"/>	20	0	0	20	40
Classroom work	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	10	0	0	10	20
Autonomous troubleshooting and / or exercises	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0	0	0	90	90
Total hours E:							150
Work load in UVIGO ECTS credits:							6

\*The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.

#### Methodologies

	Description
Master Session	The professor explains main concepts and tools included in the subject contents
Classroom work	Students do practical work: resolution of problems and questions and use of computer applications on real data

Autonomous troubleshooting and / or exercises      Students do computational work on their own

---

---

**Personalized attention**

---

	Description
Classroom work	Professors will be in touch with the students online (forum, discussion list, chat) to answer their questions. Professors will also help students in person if necessary.
Autonomous troubleshooting and / or exercises	Professors will be in touch with the students online (forum, discussion list, chat) to answer their questions. Professors will also help students in person if necessary.

---

---

**Assessment**

---

	Description	Qualification
Master Session	Students are expected to participate actively in the discussions maintained 5 in the classroom	
Classroom work	Students are expected to participate actively in the discussions maintained 5 in the classroom	
Autonomous troubleshooting and / or exercises	Students will write reports on their assignments	90

---

---

**Other comments and second call**

---

---

**Sources of information**

---

Gil Alterovitz (Editor), Marco Ramoni (Editor), Knowledge-Based Bioinformatics: From analysis to interpretation, WILEY, ISBN: 978-0-470-74831-2, 2010

Zheng Rong Yang, Machine learning approaches to bioinformatics, WORLD SCIENTIFIC, ISBN: 978-981-4287-30-2, 2010

Jonathan Pevsner, Bioinformatics and Functional Genomics, JOHN WILEY & SONS, Inc., ISBN: 978-0-470-08585-1, 2009

---

---

**Recommendations**

---

**Subjects that continue the syllabus**

---

Systems Biology/V02M123V01212

Structural Biology/V02M123V01211

Molecular Evolution/V02M123V01210

Computational Genomics/V02M123V01209

---

**Subjects that it is recommended to have taken before**

---

Statistical and Mathematical Methods in Bioinformatics/V02M123V01112

Prgramming for Bioinformatics/V02M123V01111

---