



POLITÉCNICA

INTERNATIONAL  
CAMPUS OF  
EXCELLENCE

COORDINATION PROCESS OF  
LEARNING ACTIVITIES  
PR/CL/001



E.T.S. de Ingenieria  
Agronomica, Alimentaria y de  
Biosistemas

# ANX-PR/CL/001-01

## LEARNING GUIDE

### SUBJECT

**203000023 - Machine Learning**

### DEGREE PROGRAMME

20BC - Master Universitario en Biología Computacional

### ACADEMIC YEAR & SEMESTER

2020/21 - Semester 1

## Index

---

### Learning guide

1. Description.....	1
2. Faculty.....	1
3. Skills and learning outcomes .....	2
4. Brief description of the subject and syllabus.....	3
5. Schedule.....	4
6. Activities and assessment criteria.....	6
7. Teaching resources.....	8
8. Other information.....	9

## 1. Description

---

### 1.1. Subject details

<b>Name of the subject</b>	203000023 - Machine Learning
<b>No of credits</b>	3 ECTS
<b>Type</b>	Optional
<b>Academic year of the programme</b>	First year
<b>Semester of tuition</b>	Semester 1
<b>Tuition period</b>	September-January
<b>Tuition languages</b>	English
<b>Degree programme</b>	20BC - Master Universitario en Biología Computacional
<b>Centre</b>	20 - E.T.S. de Ingeniería Agronómica, Alimentaria y de Biosistemas
<b>Academic year</b>	2020-21

## 2. Faculty

---

### 2.1. Faculty members with subject teaching role

<b>Name and surname</b>	<b>Office/Room</b>	<b>Email</b>	<b>Tutoring hours *</b>
Pedro Maria Larrañaga Mugica (Subject coordinator)	2208	pedro.larranaga@upm.es	Tu - 15:00 - 19:00 F - 15:00 - 17:00
Maria Concepcion Bielza Lozoya	2210	c.bielza@upm.es	Tu - 15:00 - 19:00 F - 15:00 - 17:00

\* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

## 3. Skills and learning outcomes \*

---

### 3.1. Skills to be learned

CE01 - Comprender las bases moleculares y las técnicas experimentales estándares más comunes en las investigaciones ómicas (genómica, transcriptómica, proteómica, metabolómica, interactómica, etc.).

CE04 - Utilizar diferentes bases de datos (incluidos los bigdata), conocer sus estructuras y ontologías, aplicar la estadística a su análisis, siendo capaz de utilizar herramientas de representación y visualización.

CE08 - Capacidad de integrar tecnologías y sistemas propios de la Inteligencia Artificial, con carácter generalista, y en contextos más amplios y multidisciplinares.

CE09 - Capacidad de interpretar los modelos de clasificación supervisada y no supervisada obtenidos al aplicar las técnicas de Aprendizaje Automático para un conjunto de datos.

CG02 - Familiarizarse con el trabajo y los métodos de la Biología Computacional en condiciones reales, adquiriendo la capacidad de diseñar aplicaciones/experimentos de forma independiente y describir, cuantificar, analizar y evaluar críticamente los resultados obtenidos.

CG05 - Que los estudiantes sean capaces de integrar conocimientos en el área de la Biología Computacional, de formular conclusiones, hipótesis o líneas de trabajo a partir de la información disponible, y de formarse una opinión fundamentada sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos.

### 3.2. Learning outcomes

RA6 - Ser capaz de llevar a cabo análisis de datos sofisticados por medio de software específico.

RA78 - RA67 - Ser capaz de identificar el tipo de método (supervisado, no supervisado, relación entre variables) a utilizar para modelizar un problema real de biología computacional

RA76 - RA64 - Ser capaz de seleccionar el modelo/algoritmo idóneo para cada método

RA77 - RA66 - Ser capaz de interpretar el modelo obtenido

\* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

## 4. Brief description of the subject and syllabus

---

### 4.1. Brief description of the subject

El Aprendizaje Automático transforma los datos en conocimiento y proporciona sistemas de propósito general que se adaptan a las circunstancias. Esta asignatura expone varios métodos para resolver problemas de clasificación supervisada, no supervisada, y relaciones entre las variables del sistema.

### 4.2. Syllabus

#### 1. Introducción al aprendizaje automático

##### 1.1. Introducción

##### 1.2. Teoría de la información

#### 2. Clasificación supervisada

##### 2.1. Métodos de evaluación

2.2. Clasificadores no probabilísticos: Vecinos más cercanos, árboles de clasificación, inducción de reglas, redes neuronales, máquinas de vectores soporte

2.3. Clasificadores probabilísticos: Análisis discriminante, regresión logística, clasificadores Bayesianos

##### 2.4. Selección de variables

##### 2.5. Metaclasificadores

##### 2.6. Clasificadores multidimensionales y multi-etiqueta

#### 3. Clasificación no supervisada

##### 3.1. Clustering no probabilístico: jerárquico, particional

##### 3.2. Clustering probabilístico: modelos de mixturas finitas

## 5. Schedule

### 5.1. Subject schedule\*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Distant / On-line	Assessment activities
1			Clases teóricas del tema 1 Duration: 02:00	
2			Clases teóricas del tema 2 Duration: 02:00	
3			Clases teóricas del tema 2 Duration: 02:00	
4			Clases teóricas del tema 2 Duration: 02:00	
5			Presentación de prácticas Duration: 02:00	Trabajo 1 (temas 1, 2.1, 2.2 y 2.4) sobre análisis de un conjunto de datos  Continuous assessment Not Presential Duration: 02:00
6			Clases teóricas del tema 2 Duration: 02:00	Test 1 (temas 1, 2.1, 2.2 y 2.4)  Continuous assessment Not Presential Duration: 00:10
7			Clases teóricas del tema 2 Duration: 02:00  Tutorías en grupo Duration: 01:00	
8			Clases teóricas del tema 2 Duration: 02:00	
9			Presentación de prácticas Duration: 02:00	Trabajo 2 (temas 2.3, 2.4, 2.5 y 2.6) sobre análisis de un conjunto de datos  Continuous assessment Not Presential Duration: 02:00
10			Clases teóricas del tema 3 Duration: 02:00  Tutorías en grupo Duration: 01:00	Test 2 (temas 2.3, 2.4, 2.5 y 2.6)  Continuous assessment Not Presential Duration: 00:10

11			<p><b>Presentación de prácticas</b> Duration: 02:00</p>	<p><b>Trabajo 3 (tema 3) sobre análisis de un conjunto de datos</b></p> <p>Continuous assessment Not Presential Duration: 02:00</p> <p><b>Test 3 (tema 3)</b></p> <p>Continuous assessment Not Presential Duration: 00:10</p>
12				
13				
14				
15				
16				
17				

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

\* The schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

## 6. Activities and assessment criteria

### 6.1. Assessment activities

#### 6.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
5	Trabajo 1 (temas 1, 2.1, 2.2 y 2.4) sobre análisis de un conjunto de datos		No Presential	02:00	16.66%	4 / 10	CE08 CE04 CE01 CG05 CE09 CG02
6	Test 1 (temas 1, 2.1, 2.2 y 2.4)		No Presential	00:10	16.66%	4 / 10	CE08 CG05 CG02
9	Trabajo 2 (temas 2.3, 2.4, 2.5 y 2.6) sobre análisis de un conjunto de datos		No Presential	02:00	16.66%	4 / 10	CE08 CE04 CE01 CG05 CE09 CG02
10	Test 2 (temas 2.3, 2.4, 2.5 y 2.6)		No Presential	00:10	16.66%	4 / 10	CE08 CG05 CG02
11	Trabajo 3 (tema 3) sobre análisis de un conjunto de datos		No Presential	02:00	16.7%	4 / 10	CE08 CE04 CE01 CG05 CE09 CG02
11	Test 3 (tema 3)		No Presential	00:10	16.66%	4 / 10	CE08 CG05 CG02

#### 6.1.2. Final examination

No se ha definido la evaluación sólo por prueba final.

#### 6.1.3. Referred (re-sit) examination

No se ha definido la evaluación extraordinaria.



## 6.2. Assessment criteria

La calificación vendrá dada a partir de la evaluación de los dos tipos de pruebas, listadas a continuación, siendo el peso de la parte de los test del 50% y de los trabajos prácticos del 50%. Para poder realizar la ponderación es necesario que el alumno obtenga una nota mínima de 4 sobre 10 en cada parte. La nota de cada parte es la media de sus tres pruebas. La realización de todos los test y las prácticas es obligatoria. Las pruebas son:

\* Tres exámenes tipo test sobre los contenidos de la asignatura.

\* Tres trabajos prácticos de modelización (usando los modelos de los temas 2-3) aplicado a un conjunto de datos elegidos por el alumno. Se evaluará el trabajo escrito y su exposición oral en el aula. Se utilizará software libre tipo WEKA, R, GeNIe.

Durante el curso se detallarán las instrucciones para la realización de las entregas.

Para la convocatoria extraordinaria de Julio, las evaluaciones serán análogas.

Los derechos y deberes de los estudiantes universitarios están desarrollados en los Estatutos de la Universidad Politécnica de Madrid (BOCM de 15 de noviembre de 2010) y en el Estatuto del Estudiante Universitario (RD 1791/2010 de 30 de diciembre).

El artículo 124 (a) de los EUPM fija como deber del estudiante ..."seguir con responsabilidad y aprovechamiento el proceso de formación, adquisición de conocimientos, y aprendizaje correspondiente a su condición de universitario"... y el artículo 13 del Estatuto del Estudiante Universitario, en el punto (d) especifica también como deber del estudiante universitario "abstenerse de la utilización o cooperación en procedimientos fraudulentos en las pruebas de evaluación, en los trabajos que se realicen o en documentos oficiales de la universidad".

En el caso de que en el desarrollo de las pruebas de evaluación se aprecie el incumplimiento de los deberes como estudiante universitario, el coordinador de la asignatura podrá ponerlo en conocimiento del Director o Decano del Centro, que de acuerdo con lo establecido en el artículo 74 (n) de los Estatutos de la UPM tiene competencias para "Proponer la iniciación del procedimiento disciplinario a cualquier miembro de la Escuela o Facultad, por propia iniciativa o a instancia de la Comisión de Gobierno" al Rector, en los términos previstos en los estatutos y normas de aplicación.

## 7. Teaching resources

### 7.1. Teaching resources for the subject

Name	Type	Notes
Libro principal	Bibliography	C. Bielza, P. Larrañaga (2020). Data-Driven Computational Neuroscience. Machine Learning and Statistical Models. Cambridge University Press
Libro 1	Bibliography	E. Alpaydin. Introduction to Machine Learning. MIT Press
Libro 2	Bibliography	R. Duda, P.E. Hart, D.G. Stork. Pattern Classification. Wiley. 2001.
Libro 4	Bibliography	L. Kuncheva. Combining Pattern Classifiers. Wiley. 2004
Libro 6	Bibliography	A. Webb. Statistical Pattern Recognition. Wiley. 2002
Libro 7	Bibliography	I. Witten, E. Frank. Data Mining. Morgan Kaufmann. 2ª ed. 2005.
software	Web resource	Software WEKA: <a href="http://www.cs.waikato.ac.nz/ml/weka">http://www.cs.waikato.ac.nz/ml/weka</a>
Libro redes bayesianas	Bibliography	D. Koller and N. Friedman (2009) Probabilistic Graphical Models: Principles and Techniques. The MIT Press
Libro aprendizaje automático	Bibliography	K.P. Murphy (2012) Machine Learning: A Probabilistic Perspective. The MIT Press
Artículo de selección de variables	Bibliography	Saeys, Y., I. Inza, and P. Larrañaga (2007) A review of feature selection techniques in bioinformatics, Bioinformatics, 23, 19, 2507-2517

## 8. Other information

---

### 8.1. Other information about the subject

La asignatura se apoya en la herramienta Moodle para proporcionar información y documentación a los alumnos, así como para la asignación de enunciados y entregas de las prácticas, la realización del test de la asignatura y la comunicación de las calificaciones de los alumnos.

Las clases magistrales y las tutorías individuales y en grupo se realizarán de forma telemática a través de la herramienta institucional UPM (Moodle Blackboard Collaborate o Microsoft Teams), todas ellas, en los horarios establecidos, así como las presentaciones orales por parte de los alumnos.