



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

2021/22 - Semester 1

Index

Learning guide

1. Description.....	1
2. Faculty.....	1
3. Prior knowledge recommended to take the subject.....	2
4. Skills and learning outcomes	2
5. Brief description of the subject and syllabus.....	3
6. Schedule.....	5
7. Activities and assessment criteria.....	7
8. Teaching resources.....	9
9. Other information.....	10

1. Description

1.1. Subject details

Name of the subject	203000023 - Machine Learning
No of credits	3 ECTS
Type	Optional
Academic year of the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	20BC - Master Universitario en Biología Computacional
Centre	20 - E.T.S. De Ingeniería Agronómica, Alimentaria Y De Biosistemas
Academic year	2021-22

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Mariano Rico Almodovar	D3205	mariano.rico@upm.es	Sin horario.
Emilio Serrano Fernandez (Subject coordinator)	2201	emilio.serrano@upm.es	Tu - 15:00 - 18:00 F - 09:00 - 12:00

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

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- Nociones de lenguaje Python

4. Skills and learning outcomes *

4.1. 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.

4.2. 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

RA77 - RA66 - Ser capaz de interpretar el modelo obtenido

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

* 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.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

El aprendizaje automático (ML) es el estudio de algoritmos que permiten que los programas mejoren automáticamente a través de la experiencia. Los algoritmos de ML construyen un modelo basado en "datos de entrenamiento" para hacer predicciones o decisiones sin estar programados explícitamente para ello. Esta asignatura expone varios métodos para la construcción de estos modelos. Además, la asignatura explora el uso de estos métodos para extraer información de un conjunto de datos y transformarla en una estructura (humanamente) comprensible para su uso posterior, es decir, para transformar datos en conocimiento.

5.2. Syllabus

1. Introducción al aprendizaje automático
2. Preproceso de datos
3. Aprendizaje supervisado
 - 3.1. Clasificación
 - 3.2. Regresión
4. Aprendizaje no supervisado
 - 4.1. Análisis por agrupamiento
 - 4.2. Reducción de dimensionalidad

6. Schedule

6.1. 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 prácticas del tema 2 Duration: 02:00	
4			Clases teóricas del tema 3 Duration: 02:00	
5			Clases prácticas del tema 3 Duration: 02:00	
6			Clases teóricas del tema 3 Duration: 02:00	
7			Clases prácticas del tema 3 Duration: 02:00	
8			Clases teóricas del tema 3 Duration: 02:00	
9			Clases prácticas del tema 3 Duration: 02:00	Proyecto práctico tema 3 Continuous assessment Presential Duration: 02:00
10			Clases teóricas del tema 4 Duration: 02:00	
11			Clases prácticas del tema 4 Duration: 02:00	
12			Clases teóricas del tema 4 Duration: 01:00 Clases prácticas del tema 4 Duration: 01:00	Proyecto práctico tema 4 Continuous assessment Presential Duration: 02:00

13			Ejercicios Duration: 02:00	
14			Ejercicios Duration: 02:00	
15			Ejercicios Duration: 02:00	
16			Ejercicios Duration: 02:00	
17				Proyecto práctico temas 3 y 4 Final examination Presential Duration: 02:00

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.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
9	Proyecto práctico tema 3		Face-to-face	02:00	70%	5 / 10	CE08 CE04 CE01 CG05 CE09 CG02
12	Proyecto práctico tema 4		Face-to-face	02:00	30%	5 / 10	CE08 CE04 CE01 CG05 CE09 CG02

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Proyecto práctico temas 3 y 4		Face-to-face	02:00	100%	5 / 10	CE08 CE04 CE01 CG05 CE09 CG02

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
-------------	----------	------	----------	--------	---------------	------------------

Proyecto práctico temas 3 y 4		Face-to-face	02:00	100%	5 / 10	CE08 CE04 CE01 CG05 CE09 CG02
-------------------------------	--	--------------	-------	------	--------	--

7.2. Assessment criteria

La evaluación "continua" y la evaluación "sólo final" son mutuamente excluyentes. Los estudiantes que quieran seguir la evaluación "solo final" deben informar al coordinador por correo electrónico (emilio.serrano@upm.es) en las dos primeras semanas del curso. De lo contrario, se sigue una evaluación continua.

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.

8. Teaching resources

8.1. 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.
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

Libro 8	Bibliography	The Data Bonanza: Improving Knowledge Discovery for Science, Engineering and Business. Atkinson, Baxter, Brezany, Corcho, Galea, Hemert, Parsons, and Snelling. John Wiley & Sons Ltd., 2013. ISBN: 978-1118398647
Libro 9	Bibliography	The Fourth Paradigm: Data-intensive Scientific Discovery. Tony Hey, Stewart Tansley and Kristin Tolle. Microsoft Research, 2009. ISBN: 978-0982544204
Libro 10	Bibliography	Data Preprocessing in Data Mining. Salvador García, Julián Luengo, Francisco Herrera. Springer-Verlag, 2015. ISBN: 978-3319102467

9. Other information

9.1. 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 y la comunicación de las calificaciones de los alumnos.

Las clases magistrales y las tutorías individuales y en grupo se realizarán de forma telemática o presencial de acuerdo a las recomendaciones oficiales debido a la situación sanitaria. En caso de ser telemática se realizará a través de la herramienta institucional UPM (Zoom o Microsoft Teams), todas ellas, en los horarios establecidos, así como posibles presentaciones orales por parte de los alumnos.

Se responderán emails y se resolverán dudas, usando sistemas de teleconferencia, en los horarios de tutorías.