



POLITÉCNICA

INTERNATIONAL
CAMPUS OF
EXCELLENCE

COORDINATION PROCESS OF
LEARNING ACTIVITIES
PR/CL/001



E.T.S. de Ingenieros
Informáticos

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000897 - Machine Learning

DEGREE PROGRAMME

10BA - Master Universitario En Ciencia De Datos

ACADEMIC YEAR & SEMESTER

2021/22 - 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

Name of the subject	103000897 - Machine Learning
No of credits	4.5 ECTS
Type	Compulsory
Academic year of the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	10BA - Master Universitario en Ciencia de Datos
Centre	10 - Escuela Tecnica Superior De Ingenieros Informaticos
Academic year	2021-22

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
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

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CECD06 - Conocer y tener capacidad para aplicar métodos de minería de datos para clasificación, modelado, segmentación y predicción a partir de un conjunto de datos

CECD07 - Capacidad para desarrollar de modelos inteligentes basados en datos

CG09 - Integración del conocimiento de distintos campos de estudio

CG11 - Conocimiento y comprensión de la informática para crear modelos, así como sistemas y procesos de información complejos

CGI01 - Adquirir conocimientos científicos avanzados del campo de la informática que le permitan generar nuevas ideas dentro de una línea de investigación

3.2. Learning outcomes

RA11 - Ser capaz de abordar los aspectos formales del proyecto inicial de una investigación

RA23 - Destreza en el uso de distintos tipos de herramientas (software o metodológicas y conceptuales) necesarias para el correcto y eficaz desarrollo de software, incluyendo entornos y librerías en el contexto de ciencia de datos.

RA2 - Conocer y aplicar técnicas de reducción de dimensionalidad y modelización de datos multivariantes.

RA14 - Presentar en público los resultados de sus trabajos de investigación

RA1 - Conocer y aplicar las principales técnicas para explorar, describir y analizar datos multivariantes.

RA34 - Apply AI techniques in real world data scenarios

* 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

4. Modelos gráficos probabilísticos

4.1. Redes bayesianas: fundamentos, inferencia y aprendizaje

4.2. Redes de Markov: fundamentos, inferencia y aprendizaje

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		Clases teóricas del tema 1 Duration: 02:00	
2	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
3	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
4	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
5	Presentaciones de prácticas Duration: 02:00		Presentaciones 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 Presential Duration: 02:00
6	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
7	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
8	Clases teóricas del tema 2 Duration: 02:00		Clases teóricas del tema 2 Duration: 02:00	
9	Presentaciones de prácticas Duration: 02:00		Presentaciones 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 Presential Duration: 02:00
10	Clases teóricas del tema 3 Duration: 02:00 Tutorías en grupo Duration: 02:00		Clases teóricas del tema 3 Duration: 02:00 Tutorías en grupo Duration: 02:00	
11	Presentaciones de prácticas Duration: 02:00		Presentaciones de prácticas Duration: 02:00	Trabajo 3 (tema 3) sobre análisis de un conjunto de datos Continuous assessment Presential Duration: 02:00

12	Clases teóricas del tema 4 Duration: 02:00		Clases teóricas del tema 4 Duration: 02:00	
13	Clases teóricas del tema 4 Duration: 02:00 Tutorías en grupo Duration: 02:00		Clases teóricas del tema 4 Duration: 02:00 Tutorías en grupo Duration: 02:00	
14	Clases teóricas del tema 4 Duration: 02:00		Clases teóricas del tema 4 Duration: 02:00	
15	Presentaciones de prácticas Duration: 02:00		Presentaciones de prácticas Duration: 02:00	Trabajo 4 (tema 4) sobre análisis de un conjunto de datos Continuous assessment Presential Duration: 02:00 Test Continuous assessment Presential Duration: 00:40
16				Trabajos 1-4 sobre análisis de un conjunto de datos Final examination Presential Duration: 01:30 Test Final examination Presential Duration: 00:40
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		Face-to-face	02:00	12.5%	4 / 10	CG09 CB10 CGI01 CECD07 CG11 CECD06
9	Trabajo 2 (temas 2.3, 2.4, 2.5 y 2.6) sobre análisis de un conjunto de datos		Face-to-face	02:00	12.5%	4 / 10	CG09 CB10 CGI01 CECD07 CG11 CECD06
11	Trabajo 3 (tema 3) sobre análisis de un conjunto de datos		Face-to-face	02:00	12.5%	4 / 10	CG09 CB10 CGI01 CECD07 CG11 CECD06
15	Trabajo 4 (tema 4) sobre análisis de un conjunto de datos		Face-to-face	02:00	12.5%	4 / 10	CB10 CGI01 CECD07 CG11 CECD06 CG09
15	Test		Face-to-face	00:40	50%	4 / 10	CG11 CECD06 CB10

6.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
16	Trabajos 1-4 sobre análisis de un conjunto de datos		Face-to-face	01:30	50%	4 / 10	CG09 CB10 CGI01 CECD07 CG11 CECD06

16	Test		Face-to-face	00:40	50%	4 / 10	CB10 CG11 CECD06
----	------	--	--------------	-------	-----	--------	------------------------

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 realización de todos los test y las prácticas es obligatoria. Las pruebas son:

* Un examen tipo test sobre los contenidos de la asignatura.

* Cuatro trabajos prácticos de modelización (usando los modelos de los temas 2-4) 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.

La situación sanitaria causada por la pandemia COVID-19 obliga a restringir el aforo de las aulas y por ello se ha decidido que la docencia de este semestre sea de presencialidad mixta. Se establecerán turnos de presencialidad dentro de los grupos, de forma que cada semana un turno asistirá a clase en el aula (columna "actividad en el aula" del cronograma), mientras el resto de los turnos se conectarán a la clase en remoto (columna "tele-enseñanza"). Y cada semana será un turno diferente el que acuda al aula. Si mejoraran las condiciones sanitarias y se pudieran impartir clases presenciales con normalidad, todos los alumnos acudirán a las aulas a recibir las clases indicadas en la columna "actividad en el aula". Si, por el contrario, empeoraran las condiciones sanitarias, todos los alumnos pasarían a conectarse a las clases en remoto de la columna "tele-enseñanza". En esta situación las pruebas de evaluación continua presenciales previstas se realizarían de forma online, sin necesidad de modificar esta guía.

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: http://www.cs.waikato.ac.nz/ml/weka
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 y la comunicación de las calificaciones de los alumnos.

Se plantea docencia presencial por turnos, de forma que todas las semanas hay estudiantes en clase y también estudiantes en casa. Por ello en el cronograma aparecen duplicadas las actividades docentes en las dos columnas.

El alumno que tenga que realizar una presentación oral (cuando le toque) lo hará presencialmente en el aula. Para el test escrito todos los alumnos estarán presentes en el aula (o aulas).

La asignatura se relaciona con el ODS4, ya que se enfoca hacia la apreciación de los límites del conocimiento actual y de la aplicación práctica de la tecnología más reciente, con pensamiento crítico, creativo, rigor científico, y gestión del conocimiento.