



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

2019/20 - 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.....	5
6. Activities and assessment criteria.....	7
7. Teaching resources.....	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	2019-20

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

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

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.

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

RA34 - Apply AI techniques in real world data scenarios

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

* 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, relaciones entre las variables del sistema y estadística espacial.

4.2. Syllabus

1. Introducción al aprendizaje automático

1.1. Reconocimiento de patrones

1.2. Preproceso de datos

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. Estadística espacial

5.1. Aleatoriedad espacial completa

5.2. Tests de ajuste por simulación

5.3. Modelos comunes de procesos puntuales espaciales

5. Schedule

5.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	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	Clases teóricas del tema 2 Duration: 02:00		Tutorías en grupo Duration: 03:00	
6	Clases teóricas del tema 2 Duration: 02:00			
7	Clases teóricas del tema 2 Duration: 02:00			
8	Clases teóricas del tema 2 Duration: 02:00			
9	Clases teóricas del tema 3 Duration: 02:00			
10	Clases teóricas del tema 3 Duration: 02:00		Tutorías en grupo Duration: 03:00	
11	Clases teóricas del tema 4 Duration: 02:00			
12	Clases teóricas del tema 4 Duration: 02:00			
13	Clases teóricas del tema 4 Duration: 02:00		Tutorías en grupo Duration: 03:00	
14	Clases teóricas del tema 5 Duration: 02:00			

15	Clases teóricas del tema 5 Duration: 02:00		Tutorías en grupo Duration: 02:00	
16	Exposición oral del trabajo individual Duration: 03:00			Realización de una exposición oral del trabajo individual sobre el manejo de software con aplicación a un conjunto de datos para los temas 2-5 Continuous assessment and final examination Duration: 03:00
17	Test Duration: 03:00			Test Continuous assessment and final examination Duration: 03:00

The independent study hours are training activities during which students should spend time on individual study or individual assignments.

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 subject schedule is based on a previous theoretical planning of the subject plan and might go through experience some unexpected changes along throughout the academic year.

6. Activities and assessment criteria

6.1. Assessment activities

6.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
16	Realización de una exposición oral del trabajo individual sobre el manejo de software con aplicación a un conjunto de datos para los temas 2-5		Face-to-face	03:00	50%	4 / 10	CG09 CECD07 CG11 CECD06
17	Test		Face-to-face	03:00	50%	5 / 10	CG11 CB10 CGI01

6.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
16	Realización de una exposición oral del trabajo individual sobre el manejo de software con aplicación a un conjunto de datos para los temas 2-5		Face-to-face	03:00	50%	4 / 10	CG09 CECD07 CG11 CECD06
17	Test		Face-to-face	03:00	50%	5 / 10	CG11 CB10 CGI01

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 las dos partes, listadas a continuación, donde para poder realizar la ponderación indicada en la tabla de evaluación (50% cada prueba) es necesario que el alumno obtenga una nota mínima en cada parte. Las dos entregas son obligatorias y constan de:

* Examen tipo test sobre los contenidos de la asignatura. El alumno deberá obtener una nota mínima de 5 sobre 10.

* Trabajo práctico de modelización (usando los modelos de los temas 2-5) aplicado a un conjunto de datos elegidos por el alumno y relacionados con la biología computacional. Se evaluará el trabajo escrito y su exposición oral en el aula. Se utilizará software libre tipo WEKA, R, GeNIe. El alumno deberá obtener una nota mínima de 4 sobre 10.

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

Para la convocatoria extraordinaria de Julio, las evaluaciones serán análogas (las dos entregas) y se exigirá presencialidad en el momento de realizarlas.

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 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 base	Bibliography	C. Bielza, P. Larrañaga (2019) Data-Driven Computational Neuroscience. Cambridge University Press
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
Libro estadística espacial	Bibliography	A. Baddeley, E. Rubak, and R. Turner (2015) Spatial Point Patterns: Methodology and Applications with R. Chapman & Hall/CRC
Artículo de selección de variables	Bibliography	Saeyns, Y., I. Inza, and P. Larrañaga (2007) A review of feature selection techniques in bioinformatics, Bioinformatics, 23, 19, 2507-2517