



POLITÉCNICA

INTERNATIONAL
CAMPUS OF
EXCELLENCE

COORDINATION PROCESS OF
LEARNING ACTIVITIES
PR/CL/001



E.T.S. de Ingenieros
Informaticos

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

105000144 - Data analytics

DEGREE PROGRAMME

10MI - Grado en Matematicas e Informatica

ACADEMIC YEAR & SEMESTER

2017/18 - Semester 2

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.....	4
6. Schedule.....	5
7. Activities and assessment criteria.....	7
8. Teaching resources.....	9

1. Description

1.1. Subject details

Name of the subject	105000144 - Data analytics
No of credits	6 ECTS
Type	Optional
Academic year of the programme	Fourth year
Semester of tuition	Semester 8
Tuition period	February-June
Tuition languages	English
Degree programme	10MI - Grado en Matematicas e Informatica
Centre	Escuela Tecnica Superior de Ingenieros Informaticos
Academic year	2017-18

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Ernestina Menasalvas Ruiz (Subject coordinator)	4303	ernestina.menasalvas@upm.es	Sin horario. hablar con la profesora
Alejandro Rodriguez Gonzalez	4302	alejandro.rg@upm.es	Sin horario.

* 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

El plan de estudios Grado en Matemáticas e Informática no tiene definidas asignaturas previas recomendadas para esta asignatura.

3.2. Other recommended learning outcomes

- ? Conocimiento y aplicación de las características, funcionalidades y estructura de las bases de datos, que permitan su adecuado uso, y el diseño y el análisis e implementación de aplicaciones basadas en ellos.
- ? Conocimiento y aplicación de los principios fundamentales y técnicas básicas de los sistemas inteligentes y su aplicación práctica.
- ? Aptitud para aplicar los conocimientos sobre estadística y optimización.
- ? Conocimiento adecuado del concepto de empresa, su organización y gestión.

4. Skills and learning outcomes *

4.1. Skills to be learned

CE25 - Conocer los campos de aplicación de las matemáticas y la informática, y tener una apreciación de la necesidad de poseer unos conocimientos técnicos profundos en ciertas áreas de aplicación; apreciación del grado de esta necesidad en, por lo menos, una situación.

CE37 - Combinar la teoría y la práctica para realizar tareas informáticas.

CG01 - Capacidad de resolución de problemas aplicando conocimientos de matemáticas, ciencias e ingeniería.

CG02 - Capacidad para el aprendizaje autónomo y la actualización de conocimientos, y reconocimiento de su necesidad en las áreas de la matemática y la informática.

CG04 - Capacidad de gestión de la información.

CG05 - Capacidad de abstracción, análisis y síntesis.

CG10 - Capacidad para usar las tecnologías de la información y la comunicación.

CG12 - Capacidad para trabajar en un contexto internacional, comunicándose en lengua inglesa y adaptándose a un nuevo entorno.

4.2. Learning outcomes

RA161 - Comprender los fundamentos y aplicar un amplio y variado repertorio de algoritmos de clustering, estimación, predicción y clasificación.

RA121 - Dado un problema real elegir las herramientas matemáticas o la tecnología informática más apropiada para su solución y diseñar su desarrollo e integración, analizando la viabilidad de su solución.

RA125 - Obtención de las técnicas necesarias para la realización de un informe o memoria sobre un trabajo realizado en un entorno socio?lingüístico nacional/internacional.

RA129 - Capacitación para formar parte de un equipo de trabajo en los diferentes cargos que se le asignen.

RA160 - Ser capaz de planificar y desarrollar un proyecto de DM a través de un proceso, demostrando su competencia en un estándar incluyendo las fases de conocimiento del negocio, conocimiento de los datos, análisis exploratorio de los datos, modelado, evaluación y explotación

RA163 - Manejar aplicaciones software para realizar DM

RA162 - Ser capaz de valorar y evaluar tanto el coste de un proyecto empresarial como su retorno sobre la inversión

RA164 - Ser capaz de trabajar en un equipo para realizar un proyecto de DM

RA120 - Dado un campo de aplicación de las matemáticas o de la informática, evaluar y diseñar la solución más apropiada para resolver alguno de sus problemas, exponiendo las dificultades técnicas y los límites de la aplicación.

* 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

La asignatura introduce a los alumnos en el desarrollo de proyectos de descubrimiento de conocimiento.

Para ello se analiza el proceso de desarrollo, los tipos de problemas y las técnicas de análisis de datos

La asignatura se acompaña de un proyecto que los alumnos irán desarrollando a lo largo del semestre

5.2. Syllabus

1. Introducción. El valor de los datos: variedad, velocidad, volumen

1.1. Big Data

1.2. Data Mining

2. Procesos de Data Mining

2.1. Proceso de KDD

2.2. CRISP-DM: fases

2.3. El preproceso

3. Tipos de problemas de data mining

3.1. Clasificación

3.2. Clustering

3.3. Asociación

4. . Evaluación de los resultados

4.1. Técnicas de evaluación

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	Assessment activities
1	Tema 1 Duration: 03:00 Lecture			
2	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
3	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
4	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
5	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
6	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
7	Tema 2 Duration: 02:00 Lecture		Tema 2 Duration: 02:00 Problem-solving class	
8	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
9	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
10	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
11	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
12	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
13	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	
14	Tema 3 Duration: 02:00 Lecture		Tema 3 Duration: 02:00 Problem-solving class	proyecto Group work Continuous assessment Duration: 00:00

15	Tema 4 Duration: 02:00 Lecture		Tema 4 Duration: 02:00 Problem-solving class	
16	Tema 4 Duration: 02:00 Lecture		Tema 4 Duration: 02:00 Problem-solving class	
17				examen continua Written test Continuous assessment Duration: 00:00 examen final Written test Final examination Duration: 00:00

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.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
14	proyecto	Group work	Face-to-face	00:00	50%	5 / 10	CG01 CG02 CG04 CG05 CG10 CG12 CE25 CE37
17	examen continua	Written test	Face-to-face	00:00	50%	5 / 10	CG01 CG04 CG05 CG10 CG12 CE25 CE37

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	examen final	Written test	Face-to-face	00:00	100%	5 / 10	CG01 CG02 CG04 CG05 CG10 CG12 CE25 CE37

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
examen	Written test	Face-to-face	02:00	100%	5 / 10	CG01 CG02 CG04 CG05 CG10 CG12 CE25 CE37

7.2. Assessment criteria

La asignatura de Procesos de data mining se evaluara mediante 1 proyecto, un examen y la asistencia y participación en clase.

Los proyectos se realizarán en grupos de **hasta 3 alumnos** de entre los matriculados de la asignatura al inicio del curso.

Para poder superar la asignatura, en la **convocatoria de junio**, se establecen los siguientes requisitos:

1. Obtener un mínimo de 50 puntos sobre los 100 disponible en el cómputo global de la evaluación sumativa
2. Es OBLIGATORIO realizar el examen y completar todas las entregas de los proyectos
3. En los informes y presentaciones de proyectos y casos prácticos se debe obtener una nota mínima igual o superior al 50% de la valoración del mismo (ver tabla de valoración sumativa)
4. En los exámenes se debe obtener una nota mínima igual o superior al 50% de la valoración de mismo (ver tabla de valoración sumativa)
5. El aspecto ?Participación en clase? forma un 10% de la nota que se conformará a partir de la asistencia a clase, la participación en la misma y la respuesta correcta a las preguntas planteadas por los profesores de la asignatura, pudiéndose realizar en cualquier momento a lo largo del curso

Para poder superar la asignatura en la **convocatoria extraordinaria de julio**, se establecen los siguientes requisitos:

1. Haber entregado los informes de proyectos y casos prácticos habiendo obtenido la nota mínima.
2. Realizar las presentaciones de los casos prácticos y proyectos

3. Realizar un examen de cada uno de los módulos de la asignatura
4. Obtener un mínimo de 50 puntos sobre los 100 disponibles en el cómputo global de los dos módulos

Para aquellos alumnos que de forma extraordinaria, no puedan realizar la evaluación continua, y previa petición por escrito durante los primeros 15 días del curso, la forma de evaluación de la asignatura será la siguiente, **siendo excluyente con la evaluación continua. Estos alumnos solo realizarán el examen final.**

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Principles of Data Mining (Adaptive Computation and Machine Learning), D Hand, MIT Press, 2001.	Bibliography	
Jiawei Han, Micheline Kamber, Data Mining : Concepts and Techniques, 2nd edition, Morgan Kaufmann, ISBN 1558609016, 2006.	Bibliography	
Data Mining Techniques: Marketing, Sales and Customer Support, Michael J. A. Berry, Gordon Linoff, John Wiley & Sons, 1997.	Bibliography	
Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson Addison Wesley (May, 2005). Hardcover: 769 pages. ISBN: 0321321367	Bibliography	
Ian Witten, Eibe Frank, Mark Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition, Morgan Kaufmann, ISBN 978-0-12-374856-0, 2011.	Bibliography	

Página web de la asignatura en moodle	Web resource	
WEKA	Others	
Sala de trabajo en grupo con ordenadores	Equipment	
aula	Equipment	