ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000875 - Data Engineering

DEGREE PROGRAMME

10AZ - Master Universitario en Innovación Digital

ACADEMIC YEAR & SEMESTER

2019/20 - Semester 2

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	103000875 - Data Engineering
No of credits	3 ECTS
Type	Optional
Academic year ot the programme	First year
Semester of tuition	Semester 2
Tuition period	February-June
Tuition languages	English
Degree programme	10AZ - Master Universitario en Innovación Digital
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 *
Fco.javier Segovia Perez (Subject coordinator)	2305	javier.segovia@upm.es	M - 10:00 - 11:00 Hablar con el profesor
Ernestina Menasalvas Ruiz	4303	ernestina.menasalvas@upm.es	M - 10:00 - 11:00 hablar con la profesora

* 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

- ? 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.

- ? Design and implementation of relational databases

- SQL

4. Skills and learning outcomes *

4.1. Skills to be learned

4.2. Learning outcomes

* 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

The course is mainly dedicated to the improvement of the development of software engineering projects by means of Data Mining.

The course is very interactive, with the development of many short projects and exposition at class. Learning by doing.

5.2. Syllabus

1. Data Engineering
    1.1. Engineering and Science: Big Data and DataMining
    1.2. Data Mining Process: CRISP-DM
    1.3. Data Mining for Software Engineering
    1.4. Data: types, quality, meassures of association
2. First steps using the tool: Data handling and preparation
3. Data Mining Modeling
    3.1. Regression
    3.2. Clasification
    3.3. Clustering
    3.4. Association

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: 02:00			
2	**Tema 2** Duration: 01:00		**Tema 2** Duration: 01:00	
3	**Tema 2** Duration: 01:00		**Tema 2** Duration: 02:00	**ASSIGNMENT 1** Continuous assessment and final examination Duration: 02:00
4	**Tema 3.1** Duration: 01:00		**Tema 3.1** Duration: 02:00	
5	**Tema 3.1** Duration: 01:00		**Tema 3.1** Duration: 02:00	
6	**Tema 3.1** Duration: 01:00		**Tema 3.1** Duration: 02:00	**ASSIGNMENT 2** Continuous assessment and final examination Duration: 02:00
7	**Tema 3.2** Duration: 01:00		**Tema 3.2** Duration: 02:00	
8	**Tema 3.2** Duration: 01:00		**Tema 3.2** Duration: 02:00	
9	**Tema 3.2** Duration: 01:00		**Tema 3.2** Duration: 02:00	
10	**Tema 3.2** Duration: 01:00		**Tema 3.2** Duration: 02:00	**ASSIGNMENT 3** Continuous assessment and final examination Duration: 02:00
11	**Tema 3.3** Duration: 01:00		**Tema 3.3** Duration: 02:00	
12	**Tema 3.3** Duration: 01:00		**Tema 3.3** Duration: 02:00	

13	**Tema 3.3** Duration: 01:00		**Tema 3.3** Duration: 02:00	**ASSIGNMENT 4** Continuous assessment and final examination Duration: 02:00
14	**Tema 3.4** Duration: 01:00		**Tema 3.4** Duration: 01:00	
15	**Tema 3.4** Duration: 01:00		**Tema 3.4** Duration: 01:00	
16	**Tema 3.4** Duration: 01:00		**Tema 3.4** Duration: 01:00	**ASSIGNMENT 5** Continuous assessment and final examination Duration: 02:00
17				**FINAL PROJECT** Continuous assessment and final examination Duration: 02:00

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 theorical planning of the subject plan and might go to 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
3	ASSIGNMENT 1		Face-to-face	02:00	10%	5 / 10	
6	ASSIGNMENT 2		Face-to-face	02:00	10%	5 / 10	
10	ASSIGNMENT 3		Face-to-face	02:00	10%	5 / 10	
13	ASSIGNMENT 4		Face-to-face	02:00	10%	5 / 10	
16	ASSIGNMENT 5		Face-to-face	02:00	10%	5 / 10	
17	FINAL PROJECT		Face-to-face	02:00	50%	5 / 10	

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
3	ASSIGNMENT 1		Face-to-face	02:00	10%	5 / 10	
6	ASSIGNMENT 2		Face-to-face	02:00	10%	5 / 10	
10	ASSIGNMENT 3		Face-to-face	02:00	10%	5 / 10	
13	ASSIGNMENT 4		Face-to-face	02:00	10%	5 / 10	
16	ASSIGNMENT 5		Face-to-face	02:00	10%	5 / 10	
17	FINAL PROJECT		Face-to-face	02:00	50%	5 / 10	

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
FINAL PROJECT AND ASSIGNMENTS		Face-to-face	02:02	100%	5 / 10	

7.2. Assessment criteria

The evaluation is based on the assignments and the final project.

Assignments and projects will be performed individually or by groups, depending on the size of the course

To pass the course it is mandatory to present all the assignments and the final project, in any modality of evaluation

Participation in class would give a 10% increase in the final score.

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, 3nd Edition, Morgan Kaufmann, ISBN 978-0-12-374856-0, 2011.	Bibliography	
Página web de la asignatura en moodle	Web resource	
IBM SPSS MODELER	Others	
Sala de trabajo en grupo con ordenadores	Equipment	
aula	Equipment	