



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros
Informaticos

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

103000935 - Data Mining (data Engineering)

PLAN DE ESTUDIOS

10BB - Eit Digital Master Programme On Fintech

CURSO ACADÉMICO Y SEMESTRE

2020/21 - Segundo semestre

Índice

Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	2
5. Descripción de la asignatura y temario.....	3
6. Cronograma.....	5
7. Actividades y criterios de evaluación.....	8
8. Recursos didácticos.....	10
9. Otra información.....	11

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	103000935 - Data Mining (data Engineering)
No de créditos	4 ECTS
Carácter	Obligatoria
Curso	Primer curso
Semestre	Segundo semestre
Período de impartición	Febrero-Junio
Idioma de impartición	Castellano
Titulación	10BB - Eit Digital Master Programme On Fintech
Centro responsable de la titulación	10 - Escuela Tecnica Superior de Ingenieros Informaticos
Curso académico	2020-21

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
Fco.javier Segovia Perez (Coordinador/a)	2305	javier.segovia@upm.es	L - 10:00 - 11:00 Hablar con el profesor
Ernestina Menasalvas Ruiz	4303	ernestina.menasalvas@upm.es	L - 10:00 - 11:00 hablar con la profesora

* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

3. Conocimientos previos recomendados

3.1. Asignaturas previas que se recomienda haber cursado

El plan de estudios Eit Digital Master Programme On Fintech no tiene definidas asignaturas previas recomendadas para esta asignatura.

3.2. Otros conocimientos previos recomendados para cursar la asignatura

- Artificial Intelligence

- Statistics

4. Competencias y resultados de aprendizaje

4.1. Competencias

CB09 - Que los estudiantes sepan transmitir de un modo claro y sin ambigüedades a un público especializado o no, resultados y conclusiones procedentes de la investigación científica y tecnológica o del ámbito de la innovación más avanzada, así como los fundamentos más relevantes y razones últimas que las sustentan

CE-FT01 - Capacidad para seleccionar las soluciones de almacenamiento para datos estructurados y no estructurados adecuadas en función del problema a resolver

CE-FT02 - Capacidad para aplicar técnicas para el análisis y la exploración de datos en un contexto financiero, y para la correcta comunicación de los resultados del análisis

CE-FT03 - Capacidad para seleccionar las técnicas y herramientas para la manipulación de datos de tipo financiero, incluyendo su procesamiento y visualización

CG03 - La capacidad de usar la lengua inglesa de manera competente, es decir, con capacitación para tareas complejas de trabajo y estudio.

CG07 - Capacidad de trabajar y comunicarse también en contextos internacionales.

4.2. Resultados del aprendizaje

RA17 - Being able to translate a data insight into a business decision and action.

RA16 - Being able to reframe a business question as a data question, reasoning about what data might be of assistance and how to obtain it

RA15 - Being able to understand how to effectively manage the analytical processes and use the results of these processes (models, clusters, etc.) as the basis for making informed, evidence-based decisions for creating value for a company

RA18 - Being able to understand the data science's implications for management and decision making in a datarich environment.

5. Descripción de la asignatura y temario

5.1. Descripción de la asignatura

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, using the IBM SPSS Modeler tool

Topics:

Data Engineering, Data Mining, Business Intelligence

CRISP-DM, or the Data Mining Process

Techniques:

- Classification

- Regression

- Association

- Clustering

5.2. Temario de la asignatura

1. INTRODUCTION TO DATA ENGINEERING
2. THE TOOL: IBM SPSS MODELER
3. THE PROCESS CRISP-DM
4. LINER REGRESSION
5. LOGISTIC REGRESSION
6. RFM ANALYSIS
7. DECISION TREES
8. NEURAL NETWORKS
9. CLUSTERING
10. NEAREST NEIGHBOR
11. ASSOCIATION RULES

6. Cronograma

6.1. Cronograma de la asignatura *

Sem	Actividad presencial en aula	Actividad presencial en laboratorio	Tele-enseñanza	Actividades de evaluación
1	INTRODUCTION TO DATA ENGINEERING Duración: 02:00 LM: Actividad del tipo Lección Magistral		INTRODUCTION TO DATA ENGINEERING Duración: 02:00 LM: Actividad del tipo Lección Magistral	
2	THE PROCESS OF DATA MINING Duración: 01:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 01:00 PR: Actividad del tipo Clase de Problemas THE PROCESS OF DATA MINING Duración: 01:00 LM: Actividad del tipo Lección Magistral	
3	Data Understanding and Data Preparation with IBM SPSS Modeler I Duración: 01:00 PR: Actividad del tipo Clase de Problemas		Data Understanding and Data Preparation with IBM SPSS Modeler I Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
4	Data Understanding and Data Preparation with IBM SPSS Modeler II Duración: 01:00 PR: Actividad del tipo Clase de Problemas		Data Understanding and Data Preparation with IBM SPSS Modeler II Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
5	LINEAR REGRESSION Duración: 01:00 LM: Actividad del tipo Lección Magistral		LINEAR REGRESSION Duración: 02:00 PR: Actividad del tipo Clase de Problemas LINEAR REGRESSION Duración: 01:00 LM: Actividad del tipo Lección Magistral	
6	Modeling with IBM SPSS Modeler I Duración: 01:00 PR: Actividad del tipo Clase de Problemas		Modeling with IBM SPSS Modeler I Duración: 02:00 PR: Actividad del tipo Clase de Problemas	ASSIGNMENT 1 TG: Técnica del tipo Trabajo en Grupo Evaluación continua y sólo prueba final No presencial Duración: 02:00
7	LOGISTIC REGRESSION Duración: 01:00 LM: Actividad del tipo Lección Magistral		LOGISTIC REGRESSION Duración: 02:00 PR: Actividad del tipo Clase de Problemas LOGISTIC REGRESSION Duración: 01:00 LM: Actividad del tipo Lección Magistral	
8	Modeling with IBM SPSS Modeler II Duración: 01:00 PR: Actividad del tipo Clase de Problemas		Modeling with IBM SPSS Modeler II Duración: 02:00 PR: Actividad del tipo Clase de Problemas	ASSIGNMENT 2 TG: Técnica del tipo Trabajo en Grupo Evaluación continua y sólo prueba final No presencial Duración: 02:00

9	DECISION TREES Duración: 01:00 LM: Actividad del tipo Lección Magistral		DECISION TREES Duración: 02:00 PR: Actividad del tipo Clase de Problemas DECISION TREES Duración: 01:00 LM: Actividad del tipo Lección Magistral	
10	RFM ANALYSIS Duración: 01:00 LM: Actividad del tipo Lección Magistral		RFM ANALYSIS Duración: 02:00 PR: Actividad del tipo Clase de Problemas RFM ANALYSIS Duración: 01:00 LM: Actividad del tipo Lección Magistral	
11	NEURAL NETWORKS Duración: 01:00 LM: Actividad del tipo Lección Magistral		NEURAL NETWORKS Duración: 02:00 PR: Actividad del tipo Clase de Problemas NEURAL NETWORKS Duración: 01:00 LM: Actividad del tipo Lección Magistral	ASSIGNMENT 3 TG: Técnica del tipo Trabajo en Grupo Evaluación continua y sólo prueba final No presencial Duración: 02:00
12	CLUSTERING Duración: 01:00 LM: Actividad del tipo Lección Magistral		CLUSTERING Duración: 02:00 PR: Actividad del tipo Clase de Problemas CLUSTERING Duración: 01:00 LM: Actividad del tipo Lección Magistral	
13	NEAREST NEIGHBOR Duración: 01:00 LM: Actividad del tipo Lección Magistral		NEAREST NEIGHBOR Duración: 02:00 PR: Actividad del tipo Clase de Problemas NEAREST NEIGHBOR Duración: 01:00 LM: Actividad del tipo Lección Magistral	
14	ASSOCIATION RULES Duración: 01:00 LM: Actividad del tipo Lección Magistral		ASSOCIATION RULES Duración: 02:00 PR: Actividad del tipo Clase de Problemas ASSOCIATION RULES Duración: 01:00 LM: Actividad del tipo Lección Magistral	ASSIGNMENT 4 TG: Técnica del tipo Trabajo en Grupo Evaluación continua y sólo prueba final No presencial Duración: 02:00
15			EXERCISES Duración: 01:00 PR: Actividad del tipo Clase de Problemas	ASSIGNMENT 5 TG: Técnica del tipo Trabajo en Grupo Evaluación continua y sólo prueba final No presencial Duración: 02:00
16				
17				FINAL PROJECT PI: Técnica del tipo Presentación Individual Evaluación continua y sólo prueba final No presencial Duración: 02:00

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

* El cronograma sigue una planificación teórica de la asignatura y puede sufrir modificaciones durante el curso

derivadas de la situación creada por la COVID-19.

7. Actividades y criterios de evaluación

7.1. Actividades de evaluación de la asignatura

7.1.1. Evaluación continua

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
6	ASSIGNMENT 1	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT02 CB09 CG03 CG07 CE-FT03 CE-FT01
8	ASSIGNMENT 2	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CB09 CG03 CG07 CE-FT03 CE-FT01 CE-FT02
11	ASSIGNMENT 3	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
14	ASSIGNMENT 4	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
15	ASSIGNMENT 5	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
17	FINAL PROJECT	PI: Técnica del tipo Presentación Individual	No Presencial	02:00	50%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03

7.1.2. Evaluación sólo prueba final

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
6	ASSIGNMENT 1	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT02 CB09 CG03 CG07 CE-FT03 CE-FT01
8	ASSIGNMENT 2	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CB09 CG03 CG07 CE-FT03 CE-FT01 CE-FT02
11	ASSIGNMENT 3	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
14	ASSIGNMENT 4	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
15	ASSIGNMENT 5	TG: Técnica del tipo Trabajo en Grupo	No Presencial	02:00	10%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03
17	FINAL PROJECT	PI: Técnica del tipo Presentación Individual	No Presencial	02:00	50%	5 / 10	CE-FT01 CE-FT02 CB09 CG03 CG07 CE-FT03

7.1.3. Evaluación convocatoria extraordinaria

Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
FINAL PROJECT AND ASSIGNMENTS	Pl: Técnica del tipo Presentación Individual	Presencial	02:02	100%	5 / 10	

7.2. Criterios de evaluación

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. Recursos didácticos

8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Principles of Data Mining (Adaptive Computation and Machine Learning), D Hand, MIT Press, 2001.	Bibliografía	
Jiawei Han, Micheline Kamber, Data Mining : Concepts and Techniques, 2nd edition, Morgan Kaufmann, ISBN 1558609016, 2006.	Bibliografía	
Data Mining Techniques: Marketing, Sales and Customer Support, Michael J. A. Berry, Gordon Linoff, John Wiley & Sons, 1997.	Bibliografía	

Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson Addison Wesley (May, 2005). Hardcover: 769 pages. ISBN: 0321321367	Bibliografía	MOST RECOMMENDED BOOK
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.	Bibliografía	
Página web de la asignatura en moodle	Recursos web	
IBM SPSS MODELER	Otros	THE TOOL WE WILL USE
Sala de trabajo en grupo con ordenadores	Equipamiento	
aula	Equipamiento	

9. Otra información

9.1. Otra información sobre la asignatura

We will use Teams for communication and collaboration