



CAMPUS
DE EXCELENCIA
INTERNACIONAL

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros
Industriales

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

53001500 - Bigdata: de los datos a las decisiones

PLAN DE ESTUDIOS

05BD - Master Universitario En Ingenieria De La Organizacion

CURSO ACADÉMICO Y SEMESTRE

2018/19 - Primer semestre

Índice

Guía de Aprendizaje

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

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	53001500 - Bigdata: de los datos a las decisiones
No de créditos	6 ECTS
Carácter	Obligatoria
Curso	Primer curso
Semestre	Primer semestre
Período de impartición	Septiembre-Enero
Idioma de impartición	Castellano
Titulación	05BD - Master universitario en ingeniería de la organización
Centro responsable de la titulación	05 - Escuela Técnica Superior de Ingenieros Industriales
Curso académico	2018-19

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
M. Camino Gonzalez Fernandez	UD Estadística	camino.gonzalez@upm.es	X - 12:30 - 14:30 V - 12:30 - 14:30
Francisco Javier Cara Cañas	UD Estadística	javier.cara@upm.es	X - 12:30 - 14:30 V - 12:30 - 14:30
Jose Manuel Mira Mcwilliams (Coordinador/a)	UD Estadística	josemanuel.mira@upm.es	X - 12:30 - 14:30 V - 12:30 - 14:30

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

3. Competencias y resultados de aprendizaje

3.1. Competencias

CB06 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB08 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

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

CG01 - Utilizar los conocimientos científicos y tecnológicos adquiridos en sus estudios de Grado en Ingeniería como recurso a integrar en la generación de soluciones a problemas de las organizaciones, sean éstos de funcionamiento o de diseño

CG02 - Analizar situaciones estructuradas y poco estructuradas de empresas y otras organizaciones, estableciendo diagnósticos apropiados, en particular, de carácter estratégico

CG03 - Concebir soluciones para afrontar problemas previamente diagnosticados, y evaluarlas desde diferentes criterios correspondientes a los distintos actores concernidos

CT01 - Aplica. Habilidad para aplicar conocimientos científicos, matemáticos y tecnológicos en sistemas relacionados con la práctica de la ingeniería

CT02 - Experimenta. Habilidad para diseñar y realizar experimentos así como analizar e interpretar datos

CT08 - Entiende los impactos. Educación amplia necesaria para entender el impacto de las soluciones ingenieriles en un contexto social global

CT09 - Se actualiza. Reconocimiento de la necesidad y la habilidad para comprometerse al aprendizaje continuo

CT10 - Conoce. Conocimiento de los temas contemporáneos

CT11 - Usa herramientas. Habilidad para usar las técnicas, destrezas y herramientas ingenieriles modernas necesarias para la práctica de la ingeniería

3.2. Resultados del aprendizaje

RA2 - Elegir y aplicar técnicas de modelado de grandes bases de datos

RA3 - Elegir y aplicar técnicas de predicción para variables cuantitativas y cualitativas

RA1 - Elegir y aplicar técnicas de análisis exploratorios de grandes bases de datos

4. Descripción de la asignatura y temario

4.1. Descripción de la asignatura

La llegada de los ordenadores y la era de la información ha ocasionado la aparición de grandes cantidades de datos en todas las áreas del conocimiento humano, desde la biología hasta astrofísica, finanzas, marketing,... Esto ha ocasionado que las decisiones que se toman en las empresas estén basadas cada vez más en las conclusiones que se extraen del análisis de dichos datos. El propósito de esta asignatura es presentar algunas de las técnicas estadísticas más relevantes para el modelado y predicción de estas grandes bases de datos, junto con aplicaciones relevantes. Los temas tratados abarcan la regresión lineal, los algoritmos de clasificación, técnicas de remuestreo, métodos de árboles, aprendizaje automático, entre otros. El enfoque de la asignatura es eminentemente práctico, incluyendo el uso de ejemplos reales y su implementación en R, que es el software estadístico de referencia en este tipo de problemas.

4.2. Temario de la asignatura

1. Introducción a R
2. Regresión lineal
3. Componentes principales
4. Análisis discriminante
5. Análisis cluster
6. Regresión logística
7. Árboles de clasificación y regresión
8. Random forest
9. Support vector machine
10. Modelos ARIMA

11. Regresión dinámica

12. Redes neuronales

5. Cronograma

5.1. Cronograma de la asignatura *

Sem	Actividad presencial en aula	Actividad presencial en laboratorio	Otra actividad presencial	Actividades de evaluación
1	<p>Introducción al Big Data Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Introducción a R Duración: 02:00 OT: Otras actividades formativas</p>			
2	<p>Regresión lineal Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Regresión lineal Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			
3	<p>Componentes principales Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Componentes principales Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			<p>Practica 1 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>
4	<p>Análisis discriminante Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Análisis discriminante Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			
5	<p>Análisis cluster Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Análisis cluster Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			<p>Practica 2 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>
6	<p>Regresión logística Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Regresión logística Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			

7	<p>Árboles de clasificación y regresión Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Árboles de clasificación y regresión Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
8	<p>Árboles de clasificación y regresión Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p> <p>Random forest Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			<p>Practica 3 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>
9	<p>Random forest Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p> <p>Support vector machine Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
10	<p>Support vector machine Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Support vector machine Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			<p>Practica 5 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>
11	<p>Modelos ARIMA Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Modelos ARIMA Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			
12	<p>Regresión dinámica Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Regresión dinámica Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			<p>Practica 6 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>
13	<p>Redes neuronales Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Redes neuronales Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
14	<p>Redes neuronales Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Redes neuronales Duración: 02:00 PR: Actividad del tipo Clase de Problemas</p>			<p>Practica 6 TI: Técnica del tipo Trabajo Individual Evaluación continua Duración: 05:00</p>

15	<p>Conclusiones Duración: 02:00 LM: Actividad del tipo Lección Magistral</p> <p>Conclusiones Duración: 02:00 LM: Actividad del tipo Lección Magistral</p>			
16				<p>Examen final EX: Técnica del tipo Examen Escrito Evaluación continua Duración: 02:00</p> <p>Examen final EX: Técnica del tipo Examen Escrito Evaluación sólo prueba final Duración: 02:00</p>
17				

Las horas de actividades formativas no presenciales son aquellas que el estudiante debe dedicar al estudio o al trabajo personal.

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.

6. Actividades y criterios de evaluación

6.1. Actividades de evaluación de la asignatura

6.1.1. Evaluación continua

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
3	Practica 1	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.67%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11
5	Practica 2	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.67%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11
8	Practica 3	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.67%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11

10	Practica 5	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.67%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11
12	Practica 6	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.67%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11
14	Practica 6	TI: Técnica del tipo Trabajo Individual	No Presencial	05:00	6.65%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11
16	Examen final	EX: Técnica del tipo Examen Escrito	Presencial	02:00	60%	0 / 10	CG02 CB08 CB10 CT01 CT02 CT08 CG01 CG03 CB06 CT09 CT10 CT11

6.1.2. Evaluación sólo prueba final

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
16	Examen final	EX: Técnica del tipo Examen Escrito	Presencial	02:00	100%	0 / 10	CG01 CG02 CB08 CB10 CT01 CT02 CT08 CG03 CB06 CT09 CT10 CT11

6.1.3. Evaluación convocatoria extraordinaria

No se ha definido la evaluación extraordinaria.

6.2. Criterios de evaluación

Se propondrán seis trabajos a lo largo del curso que supondrán el 40% de la nota final de la asignatura

7. Recursos didácticos

7.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Documentación	Bibliografía	
Análisis de datos con R	Equipamiento	
Introstatlearning	Bibliografía	James, G, Witten, D., Hastie, T. y Tibshirani, R., (2013), Introduction to statistical learning with applications in R. Springer
HastieetalESL	Bibliografía	Hastie, T., Tibshirani, R., y Friedman, J., (2009), "The Elements of statistical learning: data mining, inference and prediction", Springer
MLDummies	Bibliografía	Mueller, J.P, y Massaron, L., (2016), "machine learning for dummies". Wiley.
DeepLR	Bibliografía	Chollet, F. y Allaire, J.J., (2018), "Deep learning with R". Manning
Bishop	Bibliografía	Bishop, C., (2006), "Pattern recognition and machine learning". Springer
GoodfellowdeepL	Bibliografía	Goodfellow, I, bengio, J., y Courville, A., (2017), "Deep Learning". MIT press
MurphyML	Bibliografía	Murphy, K., (2012), "Machine learning: a probabilistic perspective". Adaptative Computation and Machine learning series.
Theodoridis	Bibliografía	Theororidis, S., y Koutroumbas, K., (2008), "Pattern Recognition". Elsevier
Peñaseries	Bibliografía	Peña, D., (2010), "Análisis de series temporales". El libro universitario

PeñaAM	Bibliografía	Peña, D., (2002), "Análisis de datos multivariantes". McGraw Hill
Peñaregresionyde	Bibliografía	Peña, D., (2010), "Regresión y diseño de experimentos". El libro universitario
Shumwaystoffer	Bibliografía	Shunway, R., y Stoffer, D., (2017), "Time series analysis and its applications with R examples". Springer
Peñafundamentos	Bibliografía	Peña, D, (2008), "Fundamentos de Estadística". El libro universitario.
Makridakis	Bibliografía	makridakis, S., Wheelwright, S. y Hyndman, R., (2007), "Forecasting: methods and applications" .
SharmaAM	Bibliografía	Sharma, S., (1995), "Applied multivariate analysis techniques"
LewisML	Bibliografía	Lewis, N., (2017), "Machine learning made easy with R. An intuitive step by step blueprint for beginners"
Geron	Bibliografía	Geron, A., (2017), "Hands-on machine learning with scikit and TensorFlow: Concepts, tools and techniques to build intelligent systems". O'Reilly
RaschkaML	Bibliografía	Raschka, S., y Mirjalili, V., (2017) , "Python machine learning"