

ANX-PR/CL/001-01
GUÍA DE APRENDIZAJE

ASIGNATURA

Data analytics

CURSO ACADÉMICO - SEMESTRE

2016-17 - Segundo semestre

Datos Descriptivos

Nombre de la Asignatura	Data analytics
Titulación	10MI - Grado en Matemáticas e Informática
Centro responsable de la titulación	Escuela Técnica Superior de Ingenieros Informáticos
Semestre/s de impartición	Octavo semestre
Materias	Optatividad
Carácter	Optativa
Código UPM	105000144
Nombre en inglés	Data analytics

Datos Generales

Créditos	6	Curso	4
Curso Académico	2016-17	Período de impartición	Febrero-Junio
Idioma de impartición	Inglés	Otros idiomas de impartición	

Requisitos Previos Obligatorios

Asignaturas Previas Requeridas

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

Otros Requisitos

El plan de estudios Grado en Matemáticas e Informática no tiene definidos otros requisitos para esta asignatura.

Conocimientos Previos

Asignaturas Previas Recomendadas

El coordinador de la asignatura no ha definido asignaturas previas recomendadas.

Otros Conocimientos Previos Recomendados

? 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.



CAMPUS
DE EXCELENCIA
INTERNACIONAL

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros Informáticos
PROCESO DE COORDINACIÓN DE LAS ENSEÑANZAS

ANX-PR/CL/001-01: GUÍA DE APRENDIZAJE



Código PR/CL/001

Competencias

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.

Resultados de Aprendizaje

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.

Profesorado

Profesorado

Nombre	Despacho	e-mail	Tutorías
Menasalvas Ruiz, Ernestina (Coordinador/a)	4303	ernestina.menasalvas@upm.es	hablar con la profesora
Fernandez Baizan, Maria Covadonga	4304	mariacovadonga.fernandez@upm.es	
Rodriguez Gonzalez, Alejandro	4302	alejandro.rg@upm.es	

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

Descripción de la Asignatura

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

Temario

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

Cronograma

Horas totales: 63 horas

Horas presenciales: 63 horas (40.4%)

Peso total de actividades de evaluación continua:
100%

Peso total de actividades de evaluación sólo prueba final:
100%

Semana	Actividad Presencial en Aula	Actividad Presencial en Laboratorio	Otra Actividad Presencial	Actividades Evaluación
Semana 1	Tema 1 Duración: 03:00 LM: Actividad del tipo Lección Magistral			
Semana 2	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 3	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 4	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 5	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 6	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 7	Tema 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 2 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 8	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	proyecto 1 Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial
Semana 9	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 10	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 11	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	

Semana 12	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 13	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 14	Tema 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 3 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 15	Tema 4 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 4 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	proyecto 1 cont Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial
Semana 16	Tema 4 Duración: 02:00 LM: Actividad del tipo Lección Magistral		Tema 4 Duración: 02:00 PR: Actividad del tipo Clase de Problemas	
Semana 17				examen continua Duración: 00:00 EX: Técnica del tipo Examen Escrito Evaluación continua Actividad presencial examen final Duración: 00:00 EX: Técnica del tipo Examen Escrito Evaluación sólo prueba final Actividad presencial

Nota.- El cronograma sigue una planificación teórica de la asignatura que puede sufrir modificaciones durante el curso.

Nota 2.- Para poder calcular correctamente la dedicación de un alumno, la duración de las actividades que se repiten en el tiempo (por ejemplo, subgrupos de prácticas") únicamente se indican la primera vez que se definen.

Actividades de Evaluación

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
8	proyecto 1	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	Sí	25%	5 / 10	CG01, CG02, CG04, CG05, CG10, CG12, CE25, CE37
15	proyecto 1 cont	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	Sí	25%	5 / 10	
17	examen continua	00:00	Evaluación continua	EX: Técnica del tipo Examen Escrito	Sí	50%	5 / 10	CG01, CG04, CG05, CG10, CG12, CE25, CE37
17	examen final	00:00	Evaluación sólo prueba final	EX: Técnica del tipo Examen Escrito	Sí	100%	5 / 10	CG01, CG02, CG04, CG05, CG10, CG12, CE25, CE37

Criterios de Evaluación

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.

Recursos Didácticos

Descripción	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	
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	
WEKA	Otros	
Sala de trabajo en grupo con ordenadores	Equipamiento	
aula	Equipamiento	