



(Procesos de Minería de Datos)

Guía de Aprendizaje – Información al estudiante

1. Datos Descriptivos

Titulación	MASTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA
Módulo	
Materia	SISTEMAS Y SERVICIOS BASADOS EN EL CONOCIMIENTO
Asignatura	PROCESOS DE MINERÍA DE DATOS
Carácter	OPTATIVO
Créditos ECTS	4
Departamento responsable	LSIIS
Especialidad	

Curso académico	2011-2012
Semestre en que se imparte	1er SEMESTRE
Idioma en el que se imparte	ESPAÑOL/INGLÉS
Página Web	http://telemaco.ls.fi.upm.es



2. Profesorado

NOMBRE Y APELLIDO	DESPACHO	Correo electrónico
Ernestina Menasalvas (Coord.)	4302	emenasalvas@fi.upm.es
Javier Segovia	4302	fsegovia@fi.upm.es
Oscar Marbán	4304	omarban@fi.upm.es

3. Conocimientos previos requeridos para poder seguir con normalidad la asignatura

Asignaturas superadas	<ul style="list-style-type: none">•
Otros resultados de aprendizaje necesarios	<ul style="list-style-type: none">• Conocimiento adecuado del concepto de empresa, su organización y gestión.• 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.



4. Objetivos de Aprendizaje

COMPETENCIAS ESPECÍFICAS ASIGNADAS A LA ASIGNATURA Y SU NIVEL DE ADQUISICIÓN		
Código	Competencia	Nivel
CE12	Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar sistemas inteligentes y sistemas basados en el conocimiento	P
CE16	Habilidad para hacer conexiones entre los deseos del consumidor o cliente y lo que la tecnología puede ofrecer	P
CE17	Capacidad para decidir entre adquirir, desarrollar o aplicar tecnología a lo largo de la amplia gama de categorías de procesos, productos y servicios de una empresa o institución	P
CE18	Capacidad para comprender el mercado, sus hábitos y necesidades de productos o servicios tecnológicos	P
CE19	Capacidad para desarrollar e implantar una solución informática en un entorno empresarial	P

Nivel de competencia: conocimiento (C), comprensión (P), aplicación (A) y análisis y síntesis (S),



RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA			
Código	Resultado de aprendizaje	Competencias asociadas	Nivel de adquisición
RA1	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	CE12, CE17	P
RA2	Comprender los fundamentos y aplicar un amplio y variado repertorio de algoritmos de clustering, estimación, predicción y clasificación.	CE12	P
RA3	Ser capaz de valorar y evaluar tanto el coste de un proyecto empresarial como su retorno sobre la inversión	CE16, CE18	P
RA4	Manejar aplicaciones software para realizar DM	CE17, CE19	P
RA5	Ser capaz de trabajar en un equipo para realizar un proyecto de DM	CE19	P



5. Sistema de evaluación de la asignatura

INDICADORES DE LOGRO		
Ref	Indicador	Relacionado con RA
I1	Realización de un plan de proyecto para un problema de data mining	RA1, RA3
I2	Con un conjunto de datos preparados resolver un ejemplo completo con especial énfasis en la fase de aplicación de algoritmos de mining	RA2, RA4
I3	Resolver en equipo cada una de las fases del proyecto con la herramienta correspondiente de data mining	RA2, RA4, RA5
I4	Exponer en equipo los resultados del proyecto	RA5
I5	Resolver en equipo cada una de las fases del proyecto buscando en la bibliografía el los últimos avances y tendencias que permitirían mejores resultados	RA1, RA5
I6	Exponer en equipo los resultados dejando de manifiesto las fortalezas y debilidades de las soluciones comparadas	RA1, RA3, RA5
I7	Explicar y razonar la necesidad y fines de las distintas fases de un proceso de DM, y como se interrelacionan y planifican con las actividades de una empresa	RA1, RA3

EVALUACION SUMATIVA			
Breve descripción de las actividades evaluables	Momento	Lugar	Peso en la calif.
Informe: Caso práctico – Plan de proyecto	Semana 3	Libre elección	15%
Presentación plan de proyecto	Semana 4	Aula	5%
Informe: Caso práctico – Análisis de datos con herramientas de data mining	Semana 13	Libre elección	25%
Presentación análisis de datos	Semana 14	Aula	5%



POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID
FACULTAD DE INFORMÁTICA
Campus de Montegancedo
Boadilla del Monte. 28660 Madrid

EVALUACION SUMATIVA			
Breve descripción de las actividades evaluables	Momento	Lugar	Peso en la calif.
Informe: Caso práctico – Replanteamiento de plan de proyecto	Semana 15	Libre elección	15%
Presentación replanteamiento plan de proyecto	Semana 16	Aula	5%
Examen	Semana 17	Aula	20%
Participación en clase	Semanas 1 a 16	Aula	10%
			Total: 100%



CRITERIOS DE CALIFICACIÓN

La asignatura de Procesos de data mining se evaluará mediante 3 proyectos, un examen y la asistencia participación en clase tal y como aparece en la tabla de evaluación sumativa.

Los proyectos se realizarán en grupos de **hasta 4 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" que aparece en la tabla sumativa 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**



CRITERIOS DE CALIFICACIÓN

continua.

1. Examen en junio en la fecha establecida en el calendario oficial de exámenes por jefatura de estudios.
2. Realización de un proyecto práctico, en grupos de hasta cuatro alumnos, compuesto por tres entregas, valorado en 20 puntos.

Para poder superar la asignatura en la **convocatoria de junio (mediante la forma extraordinaria)**, se establecen los siguientes requisitos:

1. La valoración de examen es de 80 puntos.
2. Para poder aprobar la asignatura en esta convocatoria es necesario obtener una nota mínima igual o superior al 30% de la valoración del examen.
3. En el proyecto práctico es necesario obtener una valoración mínima del 35% de la valoración del mismo.
4. Obtener un mínimo de 50 puntos sobre los 100 disponibles en el cómputo global (examen + proyecto práctico)

A los alumnos que opten por este sistema de evaluación extraordinario en la convocatoria de junio y no superen la asignatura **no se les guardará ninguna de la calificaciones** para la convocatoria de julio y sucesivas.



6. Contenidos y Actividades de Aprendizaje

CONTENIDOS ESPECÍFICOS		
Bloque / Tema / Capítulo	Apartado	Indicadores Relacionados
Tema 1: Introducción	1.1 Introducción a Data Mining	I1
Tema 2: Procesos de Data Mining	2.1 Proceso de KDD	I1
	2.2 Introducción a CRISP-DM	I1
Tema 3: Fases de CRISP-DM	3.1 Comprensión del negocio	I5
	3.2 Comprensión de los datos	I5
	3.3 Proceso	I2, I3, I5
	3.4 Modelado	I2, I3, I5
	3.5 Evaluación	I4, I5
	3.6 Despliegue	I4, I5, I6

7. Breve descripción de las modalidades organizativas utilizadas y de los métodos de enseñanza empleados

Tabla 7. Modalidades organizativas de la enseñanza








MODALIDADES ORGANIZATIVAS DE LA ENSEÑANZA		
Escenario	Modalidad	Finalidad
	Clases Teóricas	<i>Hablar a los estudiantes</i>
	Seminarios-Talleres	<i>Construir conocimiento a través de la interacción y la actividad de los estudiantes</i>
	Clases Prácticas	<i>Mostrar a los estudiantes cómo deben actuar</i>
	Prácticas Externas	<i>Completar la formación de los alumnos en un contexto profesional</i>
	Tutorías	<i>Atención personalizada a los estudiantes</i>
	Trabajo en grupo	<i>Hacer que los estudiantes aprendan entre ellos</i>
	Trabajo autónomo	<i>Desarrollar la capacidad de autoaprendizaje</i>

Tabla 5. Métodos de enseñanza

MÉTODOS DE ENSEÑANZA		
	Método	Finalidad
	Método Expositivo/Lección Magistral	Transmitir conocimientos y activar procesos cognitivos en el estudiante
	Estudio de Casos	Adquisición de aprendizajes mediante el análisis de casos reales o simulados
	Resolución de Ejercicios y Problemas	Ejercitar, ensayar y poner en práctica los conocimientos previos
	Aprendizaje Basado en Problemas (ABP)	Desarrollar aprendizajes activos a través de la resolución de problemas
	Aprendizaje orientado a Proyectos	Realización de un proyecto para la resolución de un problema, aplicando habilidades y conocimientos adquiridos
	Aprendizaje Cooperativo	Desarrollar aprendizajes activos y significativos de forma cooperativa
	Contrato de Aprendizaje	Desarrollar el aprendizaje autónomo

Se conoce como método expositivo "la presentación de un tema lógicamente estructurado con la finalidad de facilitar información organizada siguiendo criterios adecuados a la finalidad pretendida". Esta metodología -también conocida como lección (lecture)- se centra fundamentalmente en la exposición verbal por parte del profesor de los contenidos sobre la materia objeto de estudio. El término "lección magistral" se suele utilizar para denominar un tipo específico de lección impartida por un profesor en ocasiones especiales.

Análisis intensivo y completo de un hecho, problema o suceso real con la finalidad de conocerlo, interpretarlo, resolverlo, generar hipótesis, contrastar datos, reflexionar, completar conocimientos, diagnosticarlo y, en ocasiones, entrenarse en los posibles procedimientos alternativos de solución.

Situaciones en las que se solicita a los estudiantes que desarrollen las soluciones adecuadas o correctas mediante la ejercitación de rutinas, la aplicación de fórmulas o algoritmos, la aplicación de procedimientos de transformación de la información disponible y la interpretación de los resultados. Se suele utilizar como complemento de la lección magistral.

Método de enseñanza-aprendizaje cuyo punto de partida es un problema que, diseñado por el profesor, el estudiante ha de resolver para desarrollar determinadas competencias previamente definidas.

Método de enseñanza-aprendizaje en el que los estudiantes llevan a cabo la realización de un proyecto en un tiempo determinado para resolver un problema o abordar una tarea mediante la planificación, diseño y realización de una serie de actividades, y todo ello a partir del desarrollo y aplicación de aprendizajes adquiridos y del uso efectivo de recursos.

Enfoque interactivo de organización del trabajo en el aula en el cual los alumnos son responsables de su aprendizaje y del de sus compañeros en una estrategia de corresponsabilidad para alcanzar metas e incentivos grupales. Es tanto un método, a utilizar entre otros, como un enfoque global de la enseñanza, una filosofía.

Un acuerdo establecido entre el profesor y el estudiante para la consecución de unos aprendizajes a través de una propuesta de trabajo autónomo, con una supervisión por parte del profesor y durante un periodo determinado. En el contrato de aprendizaje es básico un acuerdo formalizado, una relación de contraprestación recíproca, una implicación personal y un marco temporal de ejecución.



BREVE DESCRIPCIÓN DE LAS MODALIDADES ORGANIZATIVAS UTILIZADAS Y METODOS DE ENSEÑANZA EMPLEADOS

CLASES DE TEORIA	Lección Magistral, y Estudio de Casos con la participación de empresas colaboradoras
CLASES DE PROBLEMAS	...
PRÁCTICAS	Aprendizaje basado en problemas (ABP)
TRABAJOS AUTONOMOS	Aprendizaje basado en problemas (ABP)
TRABAJOS EN GRUPO	Aprendizaje orientado a proyectos
TUTORÍAS	...



8. Recursos didácticos

RECURSOS DIDÁCTICOS	
BIBLIOGRAFÍA	•Principles of Data Mining (Adaptive Computation and Machine Learning), D Hand, MIT Press, 2001.
	•Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems), Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann, 2006.
	•Data Mining Techniques: Marketing, Sales and Customer Support, Michael J. A. Berry, Gordon Linoff, John Wiley & Sons, 1997.
	•Data Mining: Know It All Soumen Chakrabarti, Earl Cox, Eibe Frank, Ralf G ting, Jiawei Han, Xia Jiang, Micheline Kamber, Sam Lightstone, Thomas Nadeau, Richard E. Neapolitan, Dorian Pyle, Mamdouh Refaat, Markus Schneider, Toby Teorey, and Ian Witten, Morgan Kaufmann, 2008.
RECURSOS WEB	Página web de la asignatura (http://telemaco.ls.fi.upm.es)
EQUIPAMIENTO	Paquete SPSS Clementine
	Aula XXXX
	Sala de trabajo en grupo con ordenadores



9. Cronograma de trabajo de la asignatura

Semana	Actividades en Aula	Actividades en Laboratorio	Trabajo Individual	Trabajo en Grupo	Actividades de Evaluación	Otros
Semana 1 (6 horas)	<ul style="list-style-type: none"> • Introducción a Business Intelligence (2 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (1 hora) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> •
Semana 1 (7 horas)	<ul style="list-style-type: none"> • 1.1 Introducción a data mining • 2.1 Proceso de KDD • 2.2 Introducción a KDD • (2 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (3 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> •
Semana 2 (7 horas)	<ul style="list-style-type: none"> • 3.1 Comprensión del negocio • 3.2 Comprensión de los datos • (2 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (3 horas) 	<ul style="list-style-type: none"> • (I1) Informe: Caso práctico – Plan de proyecto 	<ul style="list-style-type: none"> •
Semana 2 (6 horas)	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	<ul style="list-style-type: none"> • (I1) Presentación plan de proyecto (2 horas) 	<ul style="list-style-type: none"> •
Semana 3 (6 horas)	<ul style="list-style-type: none"> • 3.3 Preproceso (2 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> •



Semana 3 (6 horas)	<ul style="list-style-type: none"> • 3.4 Modelado (2 horas) 	•	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	•	•
Semana 4 (7 horas)	<ul style="list-style-type: none"> • 3.4 Modelado (2 horas) 	•	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (3 horas) 	•	•
Semana 4 (6 horas)	•	<ul style="list-style-type: none"> • Presentación y utilización de herramienta de data mining (2 horas) 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	•	•
Semana 5 (6 horas)	•	<ul style="list-style-type: none"> • Presentación y utilización de herramienta de data mining (2 horas) 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	•	•
Semana 5 (6 horas)	•	<ul style="list-style-type: none"> • Presentación y utilización de herramienta de data mining (2 horas) 	<ul style="list-style-type: none"> • Estudio y ejercicios (2 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (2 horas) 	•	•
Semana 6 (6 horas)	<ul style="list-style-type: none"> • 3.5 Evaluación • 3.6 Despliegue (2 horas) 	•	<ul style="list-style-type: none"> • Estudio y ejercicios (1 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (3 horas) 	•	•
Semana 6 (6 horas)	•	<ul style="list-style-type: none"> • Caso práctico de evaluación y despliegue (2 horas) 	<ul style="list-style-type: none"> • Estudio y ejercicios (1 horas) 	<ul style="list-style-type: none"> • Trabajo de consolidación (3 horas) 	•	•



Semana 7 (7 horas)	•	• Caso práctico de evaluación y despliegue (2 horas)	• Estudio y ejercicios (2 horas)	• Trabajo de consolidación (3 horas)	• (12, 13) Informe: Caso práctico – Análisis de datos con herramientas de data mining	•
Semana 7 (7 horas)	•	•	• Estudio y ejercicios (2 horas)	• Trabajo de consolidación (3 horas)	• (14) Presentación análisis de datos (2 horas)	•
Semana 8 (7 horas)	•	• Caso práctico completo (2 horas)	• Estudio y ejercicios (4 horas)	• Trabajo de consolidación (3 horas)	• (15) Informe: Caso práctico – Replanteamiento de plan de proyecto	•
Semana 8 (6 horas)	•	•	• Estudio y ejercicios (2 horas)	• Trabajo de consolidación (2 horas)	• (16) Presentación replanteamiento plan de proyecto (2 horas)	•
Semana 8 (8 horas)	•	•	• Estudio y ejercicios (6 horas)	•	• (17) Examen (2 horas)	•

Nota: Para cada actividad se especifica la dedicación en horas que implica para el alumno.



POLITÉCNICA



UNIVERSIDAD POLITÉCNICA DE MADRID
FACULTAD DE INFORMÁTICA
Campus de Montegancedo
Boadilla del Monte. 28660 Madrid