



Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

## Identificación de la asignatura

<b>Asignatura</b>	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
<b>Créditos</b>	1.2 presenciales (30 horas) 3.8 no presenciales (95 horas) 5 totales (125 horas).
<b>Grupo</b>	Grupo 1, 2S(Campus Extens)
<b>Período de impartición</b>	Segundo semestre
<b>Idioma de impartición</b>	Castellano

## Profesores

Profesores	Horario de atención al alumnado					
	Hora de inicio	Hora de fin	Día	Fecha inicial	Fecha final	Despacho
Rafael Jiménez López <a href="mailto:rafa.jimenez@uib.es">rafa.jimenez@uib.es</a>	11:00h	13:00h	Jueves	24/09/2012	20/09/2013	A-213

## Titulaciones donde se imparte la asignatura

Titulación	Carácter	Curso	Estudios
Máster U. en Gestión de Recursos Humanos. Interv. Psicológica y Pedagógica	Posgrado		Posgrado
Máster U. en Gestión de Recursos Humanos. Interv. Psicológica y Pedagógica	Posgrado		Posgrado

## Contextualización

Esta asignatura semestral, de formación común y obligatoria, está enmarcada en el Plan de *Máster en Gestión de Recursos Humanos. Intervención psicológica y pedagógica* de la Universitat de les Illes Balears y pertenece al Área de Metodología de las Ciencias del Comportamiento. La asignatura consta de 5 créditos (125 horas), de los cuales, 1.2 créditos son presenciales (30 horas) y el resto no presenciales (95 horas). Las actividades presenciales se llevarán a cabo de 16 a 20 horas todos los lunes lectivos desde el 4 de marzo al 22 de abril de 2013, y aproximadamente la mitad de dichas horas se dedicará a clases teóricas y el resto a actividades prácticas. Además, se contempla también una actividad evaluativa presencial de 2 horas de duración el 20 de mayo de 2013.

La materia integrada en la asignatura pertenece al módulo de *Metodología de investigación e intervención en las organizaciones*, al que pertenecen también las asignaturas optativas *Usos y abusos de la estadística* (1r curso - 1r semestre) y *Técnicas psicométricas de validación de constructo* (1r curso - 2º semestre).

La asignatura proporciona un conjunto de resultados de aprendizaje que se concretan en los siguientes puntos:

- \* Conocer y utilizar la minería de datos para descubrir patrones y relaciones en los datos que queden recogidos en modelos de conocimiento válidos para describir la realidad bajo estudio y/o realizar predicciones.
- \* Ser capaz de realizar análisis exploratorios de los datos para identificar variables importantes en la predicción del suceso de interés, determinar la utilidad de generar nuevas variables, y obtener información





Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

que afecta a la calidad de los datos (detección de datos irrelevantes, valores omitidos, valores alejados y contradicciones).

- \* Capacidad para depurar los datos: tratamiento de valores omitidos, tratamiento de valores alejados, resolución de contradicciones.
- \* Preparación de los datos para el modelado: selección definitiva de variables, selección de registros (muestras independientes, balanceado), generación de nuevas variables predictoras, transformación de variables.
- \* Aplicación de técnicas de modelado adecuadas en función del tipo de problema formulado.
- \* Planificar y desarrollar la estrategia de obtención del modelo de conocimiento. Explorar modelos alternativos en búsqueda del que mejor resuelva el problema planteado.
- \* Conocer los fundamentos matemático-estadísticos de algunas de las principales técnicas de minería de datos.
- \* Aplicar correctamente las técnicas de minería de datos utilizando software profesional al uso, y ser capaz de interpretar adecuadamente los resultados obtenidos.
- \* Escoger los contenidos fundamentales obtenidos de la aplicación de la minería de datos para su uso y difusión en informes científicos.

Esta asignatura proporciona, por tanto, una serie de herramientas que facilitan la labor del profesional en la toma de decisiones dentro de su campo de intervención, puesto que un análisis adecuado de los datos provenientes de su campo de trabajo le permitirá detectar patrones y regularidades, y en definitiva reducir incertidumbre a la hora de adoptar una decisión sobre qué técnicas de intervención pueden ser más adecuadas en función del perfil (patrón) del objeto de intervención.

## Requisitos

La asignatura tiene como requisito esencial el conocimiento de cuestiones básicas relacionadas con la metodología de investigación y el análisis de datos: ser capaz de diseñar una investigación, ser capaz de reconocer las distintas escalas de medida, y elaborar y reunir los datos de investigación (en un paquete estadístico), y ser capaz de realizar análisis descriptivos de los datos. Como requisito recomendable, será de gran ayuda para mejorar el rendimiento en la asignatura, la capacidad de ajuste de modelos estadísticos y la capacidad de análisis estadístico inferencial (inferencia estadística de las hipótesis).

## Competencias

La asignatura tiene el propósito de contribuir a la adquisición de las competencias que se indican a continuación, las cuales forman parte del conjunto de competencias establecido en el plan de estudios del *Máster en Gestión de RRHH. Intervención psicológica y pedagógica*.

### Específicas

1. Analizar y sintetizar la información en relación con la gestión de recursos humanos, formación y dirección de equipos de trabajo (competencia a11).





Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

## Genéricas

1. Capacidad de comunicación oral y escrita (redacción de informes o documentos) (competencia c1).

## Contenidos

### Contenidos temáticos

#### Unidad Didáctica 1. Introducción a la minería de datos

- 1.1. Definición, origen y evolución de la minería de datos
- 1.2. Conceptos básicos de minería de datos
- 1.3. Ámbitos de aplicación de la minería de datos

#### Unidad Didáctica 2. Etapas en la minería de datos

- 2.1. Formulación del problema
- 2.2. Acceso a base de datos y selección de variables relevantes
- 2.3. Exploración y preprocesamiento de datos
- 2.4. Modelado predictivo/descriptivo
- 2.5. Evaluación de modelos y selección del modelo definitivo

#### Unidad Didáctica 3. Técnicas descriptivas

- 3.1. Técnicas de exploración de datos
- 3.2. Modelado descriptivo
  - 3.2.1. *Técnicas de agrupamiento*
  - 3.2.2. *Reglas de asociación*
- 3.3. Interpretación de resultados

#### Unidad Didáctica 4. Técnicas predictivas

- 4.1. Modelado predictivo
  - 4.1.1. *Redes neuronales artificiales*
  - 4.1.2. *Árboles de decisión*
  - 4.1.3. *K-nearest neighbor*
  - 4.1.4. *Naive Bayes*
  - 4.1.5. *Técnicas estadísticas clásicas*
- 4.2. Evaluación del rendimiento del modelo
- 4.3. Interpretación de resultados

## Metodología docente



Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

En este apartado se describen las actividades de trabajo presencial y no presencial (o autónomo) previstas en la asignatura.

### Actividades de trabajo presencial

Modalidad	Nombre	Tip. agr.	Descripción
Clases teóricas	Clases magistrales	Grupo grande (G)	Mediante el método expositivo, el profesor establecerá los fundamentos teóricos, así como la ejemplificación práctica de las técnicas y los procedimientos de minería de datos de las unidades didácticas que componen la materia. Además, se proporcionará información, para cada unidad didáctica, sobre el método de trabajo aconsejable y el material didáctico que deberá utilizar el alumnado para preparar de forma autónoma los contenidos. Las clases teóricas constan de 7 sesiones (sesiones de 2 a 4 horas).
Clases prácticas	Prácticas presenciales	Grupo mediano (M)	Mediante el método de resolución de ejercicios y problemas, el alumno pondrá en práctica los procedimientos y técnicas de minería de datos expuestos en las clases teóricas. Además, dichas prácticas se realizarán en aula informática, para poder hacer uso de aplicaciones apropiadas para la minería de datos. Las clases prácticas constan de 5 sesiones (sesiones de 2 a 4 horas).
Evaluación	Práctica evaluativa	Grupo mediano (M)	Se realizará una práctica evaluativa que permitirá valorar si el alumno conoce y sabe aplicar correctamente los procedimientos y técnicas de minería de datos que forman parte de la materia.

### Actividades de trabajo no presencial

Modalidad	Nombre	Descripción
Estudio y trabajo autónomo individual	Preparación de las unidades didácticas	Después de la exposición por parte del profesor en las clases magistrales, el alumno deberá profundizar en la materia. Para facilitar esta tarea, se indicará, para cada unidad didáctica, los manuales que se han de consultar.
Estudio y trabajo autónomo individual	Resolución de las prácticas individuales	Se propondrán una serie de prácticas de trabajo autónomo a lo largo del curso, consistentes en la resolución de una serie de ejercicios. Las soluciones de los ejercicios propuestos se proporcionarán en las clases prácticas presenciales.
Estudio y trabajo autónomo en grupo	Informe de prácticas: minería de datos sobre una matriz	Para desarrollar esta actividad, los estudiantes deberán organizarse en grupos de 2 ó 3 personas. A partir de una matriz de datos proporcionada por el profesor, el grupo deberá aplicar de forma práctica los procedimientos y técnicas de minería de datos desarrollados en las sesiones teóricas y prácticas. El grupo deberá entregar, en formato electrónico, un informe del trabajo desarrollado.



Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

## Riesgos específicos y medidas de protección

Las actividades de aprendizaje de esta asignatura no conllevan riesgos específicos para la seguridad y salud del alumnado y, por tanto, no es necesario adoptar medidas de protección especiales.

## Estimación del volumen de trabajo

En la siguiente tabla se presenta la distribución de horas según las diferentes actividades de trabajo presencial y de trabajo no presencial (o autónomo) planificado.

Modalidad	Nombre	Horas	ECTS	%
<b>Actividades de trabajo presencial</b>		<b>30</b>	<b>1.2</b>	<b>24</b>
Clases teóricas	Clases magistrales	14	0.56	11.2
Clases prácticas	Prácticas presenciales	14	0.56	11.2
Evaluación	Práctica evaluativa	2	0.08	1.6
<b>Actividades de trabajo no presencial</b>		<b>95</b>	<b>3.8</b>	<b>76</b>
Estudio y trabajo autónomo individual	Preparación de las unidades didácticas	25	1	20
Estudio y trabajo autónomo individual	Resolución de las prácticas individuales	25	1	20
Estudio y trabajo autónomo en grupo	Informe de prácticas: minería de datos sobre una matriz	45	1.8	36
<b>Total</b>		<b>125</b>	<b>5</b>	<b>100</b>

Al inicio del semestre estará a disposición de los estudiantes el cronograma de la asignatura a través de la plataforma UIBdigital. Este cronograma incluirá al menos las fechas en las que se realizarán las pruebas de evaluación continua y las fechas de entrega de los trabajos. Asimismo, el profesor o la profesora informará a los estudiantes si el plan de trabajo de la asignatura se realizará a través del cronograma o mediante otra vía, incluida la plataforma Campus Extens.

## Evaluación del aprendizaje del estudiante

Los resultados de aprendizaje plantificados serán valorados mediante la aplicación de una serie de procedimientos de evaluación. En la tabla del presente apartado se describe para cada procedimiento de evaluación los criterios de evaluación y su peso en la calificación de la asignatura según el itinerario evaluativo.

El estudiante obtendrá una calificación numérica entre 0 y 10 puntos para cada actividad evaluativa, la cual será ponderada según su peso a fin de obtener la calificación global de la asignatura. Para superar la asignatura, el alumno ha de obtener un mínimo de 5 puntos sobre 10 mediante la suma ponderada de todas las actividades realizadas.

No se computará la nota final cuando sólo se disponga de un elemento de evaluación de las dos actividades fijadas (práctica evaluativa, e informe de prácticas). En este caso, la nota final será de 4, sin posibilidad de recuperación.



Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

### Práctica evaluativa

Modalidad	Evaluación
Técnica	Otros procedimientos ( <b>No recuperable</b> )
Descripción	Se realizará una práctica evaluativa que permitirá valorar si el alumno conoce y sabe aplicar correctamente los procedimientos y técnicas de minería de datos que forman parte de la materia.
Criterios de evaluación	Adecuación de los procedimientos aplicados para resolver los ejercicios propuestos y exactitud de los resultados obtenidos. Adecuación de las interpretaciones y conclusiones establecidas en función de los resultados obtenidos. Dominio de los conceptos consustanciales a los contenidos de la asignatura. La práctica evaluativa tendrá una duración máxima de 2 horas, y se realizará en aula informática para poder hacer uso de aplicaciones apropiadas para la minería de datos. El criterio numérico de evaluación se adjuntará con el enunciado de la prueba.

Porcentaje de la calificación final: 50% para el itinerario A

### Informe de prácticas: minería de datos sobre una matriz

Modalidad	Estudio y trabajo autónomo en grupo
Técnica	Informes o memorias de prácticas ( <b>No recuperable</b> )
Descripción	Para desarrollar esta actividad, los estudiantes deberán organizarse en grupos de 2 ó 3 personas. A partir de una matriz de datos proporcionada por el profesor, el grupo deberá aplicar de forma práctica los procedimientos y técnicas de minería de datos desarrollados en las sesiones teóricas y prácticas. El grupo deberá entregar, en formato electrónico, un informe del trabajo desarrollado.
Criterios de evaluación	Adecuación de los procedimientos aplicados para resolver los ejercicios propuestos y exactitud de los resultados obtenidos. Adecuación de las interpretaciones y conclusiones establecidas en función de los resultados obtenidos. Dominio de los conceptos consustanciales a los contenidos de la asignatura.

Porcentaje de la calificación final: 50% para el itinerario A

### Recursos, bibliografía y documentación complementaria

#### Bibliografía básica

- \* Hernández, J., Ramírez, M.J. y Ferri, C. (2004). *Introducción a la Minería de Datos*. Madrid: Pearson Educación, S.A.
- \* Jiménez, R. (2010). *Introducción a la minería de datos (data mining): extracción de información oculta en los datos*. Manual del curso.
- \* Larose, D.T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken, NJ: Wiley.

#### Bibliografía complementaria

- \* Bigus, J.P. (1996). *Data mining with neural networks: solving business problems from application development to decision support*. New York: McGraw-Hill.
- \* Han, J. y Kamber, M. (2000). *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann.
- \* Hand, D., Mannila, H. y Smyth, P. (2001). *Principles of Data Mining*. Cambridge, MA: The MIT Press.
- \* Kantardzic, M. (2003). *Data Mining: Concepts, Models, Methods, and Algorithms*. New York: Wiley.





---

Año académico	2012-13
Asignatura	10721 - Introducción a la Minería de Datos. Extracción de Información Oculta en los
Grupo	Grupo 1, 2S
Guía docente	A
Idioma	Castellano

- \* Palmer, A. y Jiménez, R. (1999). *Introducción al Sistema Estadístico SPSS para Windows*. Palma de Mallorca: Servei de Publicacions i Intercanvi Científic. Col·lecció Materials Didàctics, 61. Sèrie de Metodologia. Universitat de les Illes Balears.
- \* Two Crows Corporation (1999). *Introduction to Data Mining and Knowledge Discovery* (3th. ed.). Maryland: Two Crows Corporation.
- \* Ye, N. (Ed.) (2003). *The Handbook of Data Mining*. Mahwah, NJ: Lawrence Erlbaum Associates.

### **Otros recursos**

---

Se proporcionará al estudiante un compendio de recursos telemáticos de interés para su formación (aplicaciones informáticas, documentos electrónicos, tutoriales, bases de datos, etc.). Estos recursos suponen un complemento importante para el aprendizaje de los contenidos de la asignatura.

