

## Proyecto docente de la asignatura

<b>Asignatura</b>	Minería de Datos		
<b>Materia</b>	Computación		
<b>Módulo</b>	Tecnologías Específicas		
<b>Titulación</b>	<b>PROGRAMA DE ESTUDIOS CONJUNTO DE GRADO EN ESTADÍSTICA Y GRADO EN INGENIERÍA INFORMÁTICA</b>		
<b>Plan</b>	551	<b>Código</b>	46970
<b>Periodo de impartición</b>	1er Cuatrimestre	<b>Tipo/Carácter</b>	Obligatoria
<b>Nivel/Ciclo</b>	Grado	<b>Curso</b>	5º
<b>Créditos ECTS</b>	6		
<b>Lengua en que se imparte</b>	Español		
<b>Profesor/es responsable/s</b>	Carlos J. Alonso González Teodoro Calonge Cano		
<b>Departamento(s)</b>	Informática (ATC, CCIA, LSI)		
<b>Datos de contacto (E-mail, teléfono...)</b>	<a href="mailto:calonso@infor.uva.es">calonso@infor.uva.es</a> , 983 185602 <a href="mailto:teodoro@infor.uva.es">teodoro@infor.uva.es</a> 983 185603		

## **1. Situación / Sentido de la Asignatura**

---

### **1.1 Contextualización**

---

La asignatura de Minería de Datos está programada como asignatura obligatoria en el Semestre 1 del 4º curso de la titulación del Grado en Ingeniería Informática, mención Computación.

La minería de datos es un término relativamente moderno que integra numerosas técnicas de análisis de datos y de extracción de modelos. El objetivo es ser capaces de extraer patrones, de describir tendencias y regularidades, de predecir comportamientos y, en general, de sacar partido a la información computerizada que nos rodea hoy en día, generalmente heterogénea y en grandes cantidades, permite a los individuos y a las organizaciones comprender y modelar de una manera más eficiente y precisa el contexto en el que deben actuar y tomar decisiones.

### **1.2 Relación con otras materias**

---

Esta asignatura, tal y como está definida, tiene una relación clara con Técnicas de Aprendizaje Automático, y se asumirá que los alumnos han adquirido las competencias de esa asignatura.

De manera indirecta también tiene relación con Estructuras de Datos, Estadística, Bases de Datos y Programación.

### **1.3 Prerrequisitos**

---

- Conocimientos de las técnicas de Aprendizaje Automático estudiadas en la asignatura de Técnicas de Aprendizaje Automático. En su defecto, conocimiento de los métodos básicos de inducción de clasificadores.
- Programación – Pues el alumno necesitará crear sus propios programas o funciones para manipular los datos, u obtener resultados.
- Estadística – Como herramienta para poder manipular y condensar los rasgos más relevantes de un conjunto de datos de tamaño grande, ya sea por el número de muestras, o por la dimensionalidad de dichas muestras.

## 2. Competencias

---

### 2.1 Generales

---

Código	Descripción
CT3	Comunicación oral y escrita en la lengua propia
	Resolución de problemas
G10	Toma de decisiones
G16	Capacidad de aplicar los conocimientos en la práctica
G21	Habilidad para trabajar de forma autónoma

### 2.2 Específicas

---

Código	Descripción
CO5	Capacidad para adquirir, obtener, formalizar y representar el conocimiento humano en una forma computable para la resolución de problemas mediante un sistema informático en cualquier ámbito de aplicación, particularmente los relacionados con aspectos de computación, percepción y actuación en ambientes o entornos inteligentes.
CO6	Capacidad para desarrollar y evaluar sistemas interactivos y de presentación de información compleja y su aplicación a la resolución de problemas de diseño de interacción persona computadora.
CO7	Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.

### 3. Objetivos

Código	Descripción
CO5.1	Comprender el ámbito de aplicación de las técnicas de minería de datos para la extracción y representación de conocimiento basada en datos.
CO7.1	Saber aplicar técnicas de selección, preparación y visualización de datos a problemas concretos y en dominios de interés práctico.
CO6.1	Emplear adecuadamente técnicas de evaluación y representación de resultados.
CO7.2	Ampliar las técnicas de aprendizaje automático conocidas y saber aplicarlas a problemas de minería de datos.
CO7.3	Usar eficazmente plataformas computacionales para resolver problemas concretos de minería de datos en áreas de interés práctico.

### 4. Contenidos

#### Bloque 1: Introducción a la Minería de Datos

Carga de trabajo en créditos ECTS:

##### a. Contextualización y justificación

Introducir el proceso de minería de datos identificando sus etapas principales.

##### b. Objetivos de aprendizaje

Comprender el ámbito de aplicación de las técnicas de minería de datos para la extracción y representación de conocimiento basada en datos.

##### c. Contenidos

Tema 1: Introducción. Etapas del proceso de minería de datos.

##### d. Métodos docentes

Ver apartado 5

##### e. Plan de trabajo

Presentación en aula y trabajo individual

##### f. Evaluación

Ver apartado 7.

##### g. Bibliografía básica

H. Witten, E. Frank and M. A. Hall, C. J. Pall. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 4th edition, 2017.

Basilio Sierra Araujo. Aprendizaje Automático: conceptos básicos y avanzados. Pearson Prentice Hall, 2006.

---

#### **h. Bibliografía complementaria**

J. Han, M. Kamber and J. Pei. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd edition, 2012.

---

#### **i. Recursos necesarios**

Proyector de video en aula, notas de la asignatura y recursos bibliográficos.

---

#### **j. Temporalización**

Semana 1

---

### **Bloque 2: Preprocesamiento y Evaluación**

Carga de trabajo en créditos ECTS:

---

#### **a. Contextualización y justificación**

Se presentan aquí las técnicas básicas de las etapas iniciales de la ingeniería de datos: preparación de datos y selección de características.

Conocer los métodos principales de evaluación de clasificadores en varios contextos.

---

#### **b. Objetivos de aprendizaje**

Saber aplicar técnicas de selección, preparación y visualización de datos a problemas concretos y en dominios de interés práctico.

Emplear adecuadamente técnicas de evaluación y representación de resultados.

Usar eficazmente plataformas computacionales para resolver problemas concretos de minería de datos en áreas de interés práctico.

---

#### **c. Contenidos**

**Tema 2: Preparación y selección de datos.**

**Tema 3: Selección de características y reducción de la dimensionalidad**

**Tema 3: Evaluación de resultados.**

#### **d. Métodos docentes**

---

Ver apartado 5

#### **e. Plan de trabajo**

---

Presentación en aula y trabajo individual.

Realización de prácticas individuales y en grupo

#### **f. Evaluación**

---

Ver apartado 7.

#### **g. Bibliografía básica**

---

H. Witten, E. Frank and M. A. Hall, C. J. Pall. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 4th edition, 2017.

Tom M. Mitchell. Machine Learning. McGraw-Hill, 1997.

#### **h. Bibliografía complementaria**

---

J. Han, M. Kamber and J. Pei. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd edition, 2012.

Ethem Alpaydin. Introduction to Machine Learning. Third edition. The MIT Press, 2014

Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to Data Mining. Addison Wesley, 2006.

Arlot, S. and Celisse, A. A survey of cross-validation procedures for model selection. Statistics Surveys, 4, 40-79, 2010. DOI:10.1214/09-SS054. Disponible en: [http://www.di.ens.fr/willow/pdfs/2010\\_Arlot\\_Celisse\\_SS.pdf](http://www.di.ens.fr/willow/pdfs/2010_Arlot_Celisse_SS.pdf)

David Page Evaluating Machine Learning Methods. <http://pages.cs.wisc.edu/~dpage/cs760/evaluating.pdf>.  
Último acceso Marzo 2019.

#### **i. Recursos necesarios**

---

Proyector de video en aula, notas de la asignatura y recursos bibliográficos.

Laboratorio de informática con software libre: Weka y Python fundamentalmente.

#### **j. Temporalización**

---

Semanas 2-6

### **Bloque 3: Extracción de conocimiento**

---

Carga de trabajo en créditos ECTS:

#### **a. Contextualización y justificación**

---

Se presentan aquí técnicas avanzadas de minería de datos que no se han estudiado en cursos anteriores.

#### **b. Objetivos de aprendizaje**

---

Ampliar las técnicas de aprendizaje automático conocidas y saber aplicarlas a problemas de minería de datos.  
Usar eficazmente plataformas computacionales para resolver problemas concretos de minería de datos en áreas de interés práctico.

#### **c. Contenidos**

---

**Tema 4: Clustering.**

**Tema 5: Multiclasificadores.**

**Tema 6: Redes recurrentes.**

**Tema 6: Reglas de asociación.**

#### **d. Métodos docentes**

---

Ver apartado 5

#### **e. Plan de trabajo**

---

Presentación en aula y trabajo individual.  
Realización de prácticas individuales y en grupo

#### **f. Evaluación**

---

Ver apartado 7.

#### **g. Bibliografía básica**

---

H. Witten, E. Frank and M. A. Hall, C. J. Pall. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 4th edition, 2017.

M. Kantardzic. Data Mining: Concepts, Models, Methods and Algorithms, 2nd Edition. Wiley 2011.

L. Kuncheva, Combining pattern classifiers, Second edition. Wiley, 2014.

#### **h. Bibliografía complementaria**

---

Christopher M. Bishop. Pattern Recognition and Machine Learning. Springer, 2006

Tan, Steinbach, Kumar - Introduction to Data Mining, 2004

J. Han, M. Kamber and J. Pei. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd edition, 2012.

Ethem Alpaydin. Introduction to Machine Learning. Third edition. The MIT Press, 2014

Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to Data Mining. Addison Wesley, 2006.

L. Breiman. Bagging predictors. Machine Learning, 24(2):123–140, 1996.

L. Rokach. Pattern classification using ensemble methods. Word Scientific Publishing, 2010.

D. Wolpert. Stacked generalization. Neural networks, 5:241–260, 1992.

#### **i. Recursos necesarios**

---

Proyector de video en aula, notas de la asignatura y recursos bibliográficos.

Laboratorio de informática con software libre: Weka y Python fundamentalmente.

#### **j. Temporalización**

---

Semanas 7-15

<b>CARGA ECTS</b>	<b>PERIODO PREVISTO DE DESARROLLO</b>
Bloque 1: 0,4	Semana 1
Bloque 2: 2.0	Semana 2-6
Bloque 3: 2.6	Semana 7-15

### **5. Métodos docentes y principios metodológicos**

---

<b>Actividad</b>	<b>Metodología</b>
<b>Clase de teoría</b>	<ul style="list-style-type: none"><li>● Clase magistral participativa</li><li>● Estudio de casos en aula</li><li>● Resolución de problemas</li></ul>

<b>Clase práctica</b>	<ul style="list-style-type: none"><li>• Clase magistral participativa</li><li>• Realización de prácticas, cerradas y abiertas. <b>Se proporcionarán guiones con soluciones propuestas tras las entregas.</b></li></ul>
<b>Seminarios</b>	<ul style="list-style-type: none"><li>• Talleres de aprendizaje</li></ul>
<b>Tutoría</b>	<ul style="list-style-type: none"><li>• Discusión de los contenidos teóricos y de las prácticas</li></ul>



## 6. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Clases teórico-prácticas (T/M)	30	Estudio y trabajo autónomo individual	30
Clases prácticas de aula (A)		Estudio y trabajo autónomo grupal	60
Laboratorios (L)	20		
Prácticas externas, clínicas o de campo			
Seminarios (S)	6		
Tutorías grupales (TG)			
Evaluación	4		
<b>Total Presencial</b>	<b>60</b>	<b>Total no Presencial</b>	<b>90</b>

## 7. Sistema y características de la evaluación

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Examen práctico	40%	A realizar en el laboratorio durante la misma jornada que el examen final escrito
Examen final escrito	60%	Periodo de exámenes

### CRITERIOS DE CALIFICACIÓN

- **Convocatoria ordinaria:**
  - La nota final será el resultado de la suma ponderada del examen práctico y el final escrito con los porcentajes indicados en la tabla anterior. No obstante, para superar la asignatura es preciso que en ambas partes (práctico y final escrito) se obtenga un mínimo de 3 puntos sobre 10. Adicionalmente, la entrega de las prácticas, que es una actividad optativa, se valorará sobre un punto, que se añadirá a la nota final, **siempre que ésta sea igual o superior a 4.5 puntos.**
- **Convocatoria extraordinaria:**
  - Los mismos criterios que en la convocatoria ordinaria. Adicionalmente, si alguna de las partes ha sido superada en la convocatoria ordinaria del presente curso, y si el estudiante así lo desea, se le conservará la puntuación para calcular su calificación final.

## 8. Consideraciones finales

---

