

**Minería de Datos****2015/2016**

Código: 43347  
 Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
4314660 Ingeniería Informática / Computer Engineering	OT	2	1

**Contacto**

Nombre: Carlos Alejandro Parraga

Correo electrónico: CarlosAlejandro.Parraga@uab.cat

**Uso de idiomas**

Lengua vehicular mayoritaria: español (spa)

**Otras observaciones sobre los idiomas**

La comprensión de literatura científica en lengua Inglesa es importante para seguir este curso, por lo tanto se requiere un mínimo conocimiento de inglés escrito.

**Prerequisitos**

- Estar motivado para aprender algo de Python.
- Ganas de trabajar con información de redes sociales.
- Conocimientos mínimos de análisis de datos, reconocimiento de patrones y aprendizaje computacional.

**Objetivos y contextualización**

Este módulo se inicia explicando qué es la minería de datos haciendo, un análisis de cómo la minería de datos es tratada en las empresas por las distintas disciplinas que contribuyen a este campo. Se plantea una arquitectura básica con la capacidad de extraer datos y de presentar los resultados de su análisis de una manera convincente y útil. Seguidamente se explican las técnicas de modelización más utilizadas en grandes conjuntos (incluso masivos como en la Web) de datos. La comparativa se hará teniendo en cuenta medidas básicas de evaluación del rendimiento de estos sistemas.

**Competencias**

- Capacidad para analizar las necesidades de información que se plantean en un entorno y llevar a cabo en todas sus etapas el proceso de construcción de un sistema de información.
- Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento.
- Capacidad para asegurar, gestionar, auditar y certificar la calidad de los desarrollos, procesos, sistemas, servicios, aplicaciones y productos informáticos.
- Capacidad para dirigir, planificar y supervisar equipos multidisciplinares.
- Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática.
- Capacidad para la aplicación de los conocimientos adquiridos y de resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinares, siendo capaces de integrar estos conocimientos.
- Capacidad para la dirección de obras e instalaciones de sistemas informáticos, cumpliendo la normativa vigente y asegurando la calidad del servicio.

- Capacidad para la dirección general, dirección técnica y dirección de proyectos de investigación, desarrollo e innovación, en empresas y centros tecnológicos, en el ámbito de la Ingeniería Informática.
- Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la ingeniería informática.
- Concretar e indicar resultados asegurando altos niveles de rendimiento y calidad.
- Demostrar un espíritu emprendedor e innovador en la búsqueda de nuevos espacios/ámbitos en su campo de trabajo, con una visión amplia de las posibilidades de la carrera profesional en el ámbito de la Ingeniería en Informática.
- 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
- 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
- Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
- Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

## Resultados de aprendizaje

1. Capacidad para dirigir, planificar y supervisar equipos multidisciplinares.
2. Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática.
3. Capacidad para la aplicación de los conocimientos adquiridos y de resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinares, siendo capaces de integrar estos conocimientos.
4. Capacidad para la dirección de obras e instalaciones de sistemas informáticos, cumpliendo la normativa vigente y asegurando la calidad del servicio.
5. Capacidad para la dirección general, dirección técnica y dirección de proyectos de investigación, desarrollo e innovación, en empresas y centros tecnológicos, en el ámbito de la Ingeniería Informática.
6. Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la ingeniería informática.
7. Conceptualizar alternativas de soluciones complejas a problemas de minería de datos y crear prototipos que demuestren la validez del sistema propuesto.
8. Concretar e indicar resultados asegurando altos niveles de rendimiento y calidad.
9. Demostrar un espíritu emprendedor e innovador en la búsqueda de nuevos espacios/ámbitos en su campo de trabajo, con una visión amplia de las posibilidades de la carrera profesional en el ámbito de la Ingeniería en Informática.
10. Identificar las mejores representaciones que se puedan definir para la resolución de problemas de minería de datos.
11. Identificar los problemas básicos que se deben resolver en la minería de datos, así como los algoritmos específicos.
12. Planificar, desarrollar, evaluar y gestionar una solución a un problema concreto de minería de datos.
13. 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
14. 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
15. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
16. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades
17. Seleccionar las técnicas aprendidas y entrenarlas para solucionar un proyecto concreto de minería de datos.

## Contenido

1. Introducción: de los datos al conocimiento
2. Preparación de los datos
3. Presentación del conocimiento
4. Modelización supervisada
5. Modelización no supervisada
6. Minería Web

## Metodología

- Clases expositivas
- Aprendizaje basado en problemas/proyectos
- Presentación/exposición oral de trabajos
- Debates
- Participación en actividades complementarias
- Elaboración de trabajos
- Estudio personal

## Actividades

Título	Horas	ECTS	Resultados de aprendizaje
<b>Tipo: Dirigidas</b>			
Clases presenciales	41	1,64	17, 16, 7, 8, 11, 10, 12, 6, 15, 2
<b>Tipo: Supervisadas</b>			
Trabajo Práctico	15	0,6	17, 16, 7, 8, 4, 5, 9, 11, 10, 3, 1, 12, 6, 15, 2, 13
<b>Tipo: Autónomas</b>			
Trabajo fuera de clase	90	3,6	17, 16, 7, 14, 8, 4, 5, 9, 11, 10, 3, 1, 12, 6, 15, 2, 13

## Evaluación

Detalles del Proyecto:

- Defensa oral del trabajo/s (10%)
- Participación activa en el proyecto (10%)
- Autoevaluación del equipo de trabajo (10%)
- Entrega de informe (40%)

## Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Exámenes escritos	10-25%	4	0,16	17, 16, 7, 14, 8, 11, 10, 12, 15, 13
Participación en clase y trabajo en casa	5-20%	0	0	17, 16, 7, 14, 8, 4, 5, 9, 11, 10, 3, 1, 12, 6, 15, 2, 13

---

Projecto	60-75%	0	0	17, 16, 7, 14, 8, 4, 5, 9, 11, 10, 3, 1, 12, 6, 15, 2, 13
----------	--------	---	---	---

---

## Bibliografía

Data Mining: Practical Machine Learning Tools and Techniques by Ian H. Witten & Eibe Frank & Mark A. Hall, Elsevier 2011.

Data Mining: Concepts and Techniques by Jiawei Han & Micheline Kamber & Jian Pei, Morgan Kaufmann 2001.

Mining the Social Web - Analyzing Data From Facebook, Twitter, Linkedin and Other Social Media Sites by Matthew A. Russell, O'Reilly 2011.