

**FICHA IDENTIFICATIVA****Datos de la Asignatura**

<b>Código</b>	46574
<b>Nombre</b>	Aprendizaje máquina (I)
<b>Ciclo</b>	Máster
<b>Créditos ECTS</b>	6.0
<b>Curso académico</b>	2023 - 2024

**Titulación(es)**

<b>Titulación</b>	<b>Centro</b>	<b>Curso</b>	<b>Periodo</b>
2262 - M.U. en Ciencia de Datos	Escuela Técnica Superior de Ingeniería	1	Primer cuatrimestre

**Materias**

<b>Titulación</b>	<b>Materia</b>	<b>Caracter</b>
2262 - M.U. en Ciencia de Datos	6 - Aprendizaje máquina (I)	Obligatoria

**Coordinación**

<b>Nombre</b>	<b>Departamento</b>
FERRI RABASA, FRANCESC JOSEP	240 - Informática
GOMEZ CHOVA, LUIS	242 - Ingeniería Electrónica
NAVARRO QUILES, ANA	130 - Estadística e Investigación Operativa

**RESUMEN**

Esta asignatura se basa en aprender los modelos lineales y sus extensiones, las principales arquitecturas neuronales y algoritmos de aprendizaje no supervisado. Conocer qué es una máquina de vectores soporte y su diferencia con los perceptrones multicapa. Aprender los conceptos de entropía y ganancia en información. Conocer los árboles de decisión y su construcción. Conocer el concepto de clustering y los algoritmos iniciales planteados incluyendo métodos adaptativos y autoorganizativos. Conocer las ventajas de la combinación de clasificadores/modelizadores. Conocer las extensiones de los tipos de aprendizaje máquina clásicos.



## CONOCIMIENTOS PREVIOS

### Relación con otras asignaturas de la misma titulación

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

### Otros tipos de requisitos

## COMPETENCIAS

### 2262 - M.U. en Ciencia de Datos

- Que los/las estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo
- Ser capaces de valorar la necesidad de completar su formación técnica, científica, en lenguas, en informática, en literatura, en ética, social y humana en general, y de organizar su propio autoaprendizaje con un alto grado de autonomía
- Habilidad para defender criterios con rigor y argumentos, y de exponerlos de forma adecuada y precisa
- Capacidad de acceso y gestión de la información en diferentes formatos para su posterior análisis con el fin de obtener conocimiento a partir de datos.
- Capacidad para trabajar en equipo para llegar a soluciones de problemas interdisciplinarios usando técnicas de análisis de datos.
- Ser capaces de acceder a herramientas de información (bibliográficas y de empleo) y utilizarlas apropiadamente.
- Ser capaces de asumir la responsabilidad de su propio desarrollo profesional y de su especialización en uno o más campos de estudio, aplicando los conocimientos adquiridos en la identificación de salidas profesionales y yacimientos de empleo.
- Extraer conocimiento de conjuntos de datos en diferentes formatos.
- Entender la utilidad de la ciencia de datos y sus elementos asociados, así como su aplicación en la resolución de problemas, eligiendo las técnicas más adecuadas a cada problema, aplicando de forma correcta las técnicas de evaluación y, finalmente, interpretando los modelos y resultados.
- Capacidad para resolver problemas de clasificación, modelización, segmentación y predicción a partir de un conjunto de datos.
- Modelar la dependencia entre una variable respuesta y varias variables explicativas, en conjuntos de datos complejos, mediante técnicas de aprendizaje máquina, interpretando los resultados obtenidos.



## RESULTADOS DE APRENDIZAJE

- Saber construir e interpretar modelos de regresión lineal simple y múltiple.
- Diagnóstico y validación de modelos de regresión.
- Comparación de modelos de regresión. Selección de variables.
- Aprender las principales arquitecturas neuronales, perceptrón multicapa y mapas autoorganizados (S.O.M).
- Conocer qué es una máquina de vectores soporte y su diferencia con los perceptrones multicapa.
- Aprender los conceptos de entropía y ganancia en información.
- Conocer los árboles de decisión y su construcción.
- Conocer el concepto de clustering y los algoritmos iniciales planteados.
- Conocer las ventajas de la combinación de clasificadores/modelizadores.

## DESCRIPCIÓN DE CONTENIDOS

### 1. Aprendizaje supervisado

Regresión, correlación y causalidad; Distribución normal multivariante; Regresión simple y regresión múltiple; Diagnóstico y validación de modelos de regresión múltiple; Estimación y contraste de hipótesis; Tabla ANOVA; Predicción; Comparación de modelos de regresión; Selección de variables. Perceptrón multicapa, (algoritmos de aprendizaje, preprocesado de las entradas, selección de la arquitectura, enfoque bayesiano); máquinas de vectores soporte (diseño de kernels, aprendizaje por múltiples kernels, multiclase, one-class), árboles de decisión (poda, extracción de reglas).

### 2. Combinación de expertos

Bagging, boosting, random forest, extremely randomized trees

### 3. Aprendizaje no supervisado

k-means (algoritmo EM), clustering jerárquico, elección del número de clusters, métodos basados en densidades y métodos adaptativos

**VOLUMEN DE TRABAJO**

ACTIVIDAD	Horas	% Presencial
Prácticas en laboratorio	29,00	100
Clases de teoría	27,00	100
Clases teórico-prácticas	4,00	100
Elaboración de trabajos individuales	20,00	0
Estudio y trabajo autónomo	12,00	0
Lecturas de material complementario	3,00	0
Preparación de actividades de evaluación	12,00	0
Preparación de clases de teoría	20,00	0
Preparación de clases prácticas y de problemas	13,00	0
Resolución de casos prácticos	10,00	0
<b>TOTAL</b>	<b>150,00</b>	

**METODOLOGÍA DOCENTE**

Actividades teóricas. Desarrollo expositivo de la materia con la participación del estudiante en la resolución de cuestiones puntuales. Realización de cuestionarios individuales de evaluación.

Actividades prácticas. Aprendizaje mediante resolución de problemas, ejercicios y casos de estudio a través de los cuales se adquieren competencias sobre los diferentes aspectos de la materia.

Trabajos en laboratorio y/o aula ordenador. Aprendizaje mediante la realización de actividades desarrolladas de forma individual o en grupos reducidos y llevadas a cabo en aulas de ordenador.

**EVALUACIÓN**

La evaluación del aprendizaje de los conocimientos y competencias conseguidas por los estudiantes se hará de forma continuada a lo largo del curso, y constará de los siguientes bloques de evaluación:

1. Ejercicios y trabajos entregados durante el curso y/o exámenes parciales: 40% de la nota final.
2. Examen final: 60% de la nota final.

Las calificaciones obtenidas en el apartado 1 se conservarán en las dos convocatorias del curso académico en que hayan sido realizadas, dado que su evaluación sólo es posible en el periodo de docencia.



## REFERENCIAS

### Básicas

- Richard O. Duda (2016) Pattern Classification, Third Edition, John Wiley & Sons Inc.
- Trevor Hastie, Robert Tibshirani, Jerome Friedman (2011) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, Springer (Series in Statistics).
- Christopher Bishop (2010) Pattern Recognition and Machine Learning, First Edition, Springer (Information Science and Statistics).
- Ethem Alpaydin (2014) Introduction to Machine Learning, Third Edition, The Mit Press (Adaptive Computation and Machine Learning Series).
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics)
- Peter K. Dunn, Gordon K. Smyth (2018) Generalized Linear Models with Examples in R Springer (Springer Texts in Statistics).

### Complementarias

- Sebastian Raschka (2015) Python Machine Learning, Packt Publishing