

## **FICHA DE ASIGNATURA**

**Título:** Aprendizaje no supervisado

**Descripción:**

Esta asignatura ayuda al estudiante a profundizar en el conocimiento específico de la subárea del aprendizaje automático no supervisado. Más concretamente, el alumno adquirirá una visión detallada de los aspectos teóricos del aprendizaje no supervisado en diferentes escenarios según el tipo de datos que se maneja. Tendrá una especial relevancia el análisis de agrupamientos o *clustering*. También se explorará la subárea intermedia del aprendizaje semi-supervisado, a medio camino entre el mundo supervisado y el no supervisado. Desde el punto de vista práctico se ofrecen las herramientas para su resolución. En concreto, se conocerán diferentes técnicas que se utilizan para resolver cada uno de los problemas estudiados en la asignatura.

**Carácter:** *Obligatoria*

**Créditos ECTS:** 6

**Contextualización:** En el aprendizaje automático, se modelan problemas reales mediante técnicas matemático-computacionales, y generalmente se diferencian dos tipos de problemas: los predictivos, que buscan aprender un modelo que anticipe el valor de cierta variable por el momento desconocido, y los descriptivos, que tratan de modelar la estructura del problema partiendo de las (di)similitudes entre los ejemplos disponibles. Esta asignatura explora esta segunda subárea: el aprendizaje no supervisado.

**Modalidad:** *Online*

**Temario:**

- Introducción: Minería de datos, Aprendizaje No Supervisado vs Aprendizaje Supervisado, Medidas de distancia.
- Análisis de agrupamientos o clustering: Basado en centroides (k-means, k-medoids), Jerárquico, Espectral, Basado en densidades (Mean-shift, DBSCAN), basado en distribuciones (Mixtura de gaussianas, algoritmo Expectation-Maximization).
- Otras técnicas de análisis de agrupamiento: Biclustering.
- Aprendizaje semi-supervisado: Expectation-Maximization, basado en grafos, Co-training.
- Reducción de dimensionalidad: Principal Component Analysis, Independent Component Analysis.
- Otras técnicas no supervisadas: Análisis de grafos (Algoritmo Page Rank), Reglas de asociación (Algoritmo A priori).
- Clustering profundo: Autoencoders variacionales y aprendizaje adversarial para clustering.

**Competencias Específicas:**

- ⇒ CE1. Conocer los fundamentos de la ingeniería de datos (modelado, ingesta, almacenamiento, procesado, análisis y visualización), las técnicas de rastreo, procesamiento, indexación y recuperación de información.
- ⇒ CE2. Conocer y familiarizarse con el uso de las librerías y herramientas más comunes en la industria.
- ⇒ CE5. Dominar técnicas avanzadas en el ámbito del aprendizaje automático y optimización.
- ⇒ CE6. Capacidad para la abstracción de información a partir de big data mediante algoritmos de inteligencia artificial.
- ⇒ CE8. Capacidad para aplicar metodologías de diseño, implementación y testeo de frameworks de aprendizaje.
- ⇒ CE9. Evaluar de manera equitativa diferentes soluciones basadas en inteligencia artificial y elección de la más efectiva en base a ciertos criterios.
- ⇒ CE10. Interpretar los resultados de evaluación obtenidos de cualquier modelo predictivo o algoritmo basado en inteligencia artificial.
- ⇒ CE11. Consideración del rol de la inteligencia artificial en el mundo actual.
- ⇒ CE12. Desarrollar y aplicar técnicas de Inteligencia artificial para la resolución de problemas en el mundo laboral.
- ⇒ CE15. Capacidad para seguir los avances tecnológicos en el área de la inteligencia artificial.

**Actividades Formativas:**

<b>Actividad Formativa</b>	<b>Horas</b>	<b>Presencialidad</b>
Clases expositivas	40	0%
Clases prácticas sobre laboratorio informático. Estudio de casos, resolución de problemas y diseño de proyectos	30	0%
Tutorías online	10	30%
Trabajo autónomo	70	0%

**Metodologías docentes:**

Lección magistral

Estudio de casos

Resolución de problemas

Simulaciones

Laboratorio Informático Virtual

Seguimiento

**Sistema de Evaluación:**

<b>Sistemas de evaluación</b>	<b>Ponderación mínima</b>	<b>Ponderación máxima</b>
Evaluación de portfolio. Informe sobre resolución de problemas o estudio de casos	20%	30%
Evaluación del portfolio. Informe sobre diseño de proyectos	10%	30%
Participación en grupos de debate	0%	10%
Evaluación de la prueba	40%	60%

**Bibliografía:**

Hastie, T., Tibshirani, R., Friedman, J. (2008). *The Elements of Statistical: Data Mining, Inference, and Prediction*. 2ª edición. Springer.

Aggarwal, C. C., Reddy, C. K. (Eds.) (2014). *Data clustering: algorithms and applications*. Chapman & Hall/CRC.

Witten, I. H., Frank, E., Hall, M. A., Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. 4ª edición. Morgan Kaufmann.