

GUÍA DOCENTE ABREVIADA DE LA ASIGNATURA

261 - Minería de Datos

Máster Universitario en Ciencia de Datos / Master in Data Science

Curso Académico 2023-2024

| 1. DATOS IDENTIFICATIVOS | | | | | |
|--------------------------|---|------------------|-------------------|----------------------|----------------------|
| Título/s | Máster Universitario en Ciencia de Datos / Master in Data Science | | | Tipología v Curso | Obligatoria. Curso 1 |
| Centro | Facultad de Ciencias | | | | |
| Módulo / materia | FUNDAMENTOS MÉTODOS EN CIENCIA DE DATOS | | | | |
| Código y denominación | 261 - Minería de Datos | | | | |
| Créditos ECTS | 6 | Cuatrimestre | Cuatrimestral (1) | | |
| Web | | | | | |
| Idioma de impartición | Español | English friendly | Sí | Forma de impartición | Presencial |

| | | | | | |
|----------------------|---|--|--|--|--|
| Departamento | DPTO. MATEMATICA APLICADA Y CIENCIAS DE LA COMPUTACION | | | | |
| Profesor responsable | SIXTO HERRERA GARCIA | | | | |
| E-mail | sixto.herrera@unican.es | | | | |
| Número despacho | E.T.S. de Ingenieros de Caminos, Canales y Puertos. Planta: + 1. DESPACHO PROFESORES (1034) | | | | |
| Otros profesores | STEVEN JOHAN MARIA VAN VAERENBERGH JOAQUIN BEDIA JIMENEZ RODRIGO GARCIA MANZANAS ANA CASANUEVA VICENTE | | | | |

3.1 RESULTADOS DE APRENDIZAJE

- Entender las técnicas para la descripción estadística de un conjunto de datos y saber aplicar críticamente las que mejor se adapten a un caso práctico.
- Conocer y aplicar el concepto de regularización en problemas de minería de datos.
- Saber identificar los distintos problemas de minería de datos (asociación, segmentación, clasificación y predicción) y saber qué técnicas se aplican en cada caso.
- Entender las principales metodologías para aprendizaje automático (basado en casos, supervisado o no supervisado) .
- Entender el funcionamiento y saber aplicar las principales técnicas de minería de datos.
- Conocer el concepto de generalización y entender el problema del sobreajuste.
- Saber validar de forma crítica los modelos aprendidos y saber identificar el modelo más adecuado para un caso práctico concreto.
- Saber utilizar paquetes de software para resolver casos prácticos de minería de datos y analizar críticamente las soluciones obtenidas.

4. OBJETIVOS

Esta materia tiene como objetivo proporcionar a los estudiantes conocimientos básicos sobre los métodos estadísticos para la descripción, presentación y análisis de datos, así como los principios generales para la aplicación práctica de las principales técnicas automáticas de modelización y predicción de minería de datos para el tratamiento de grandes masas de datos.

6. ORGANIZACIÓN DOCENTE

| CONTENIDOS | |
|------------|--|
| 1 | Introducción y perspectiva histórica; Problemas Canónicos, Paradigmas de Aprendizaje y Validación; Aprendizaje no supervisado con Reglas de Asociación; Aprendizaje supervisado con técnicas de vecinos k-nn; Cross-Validación y Aprendizaje |
| 2 | Regularización y selección de variables Reducción de la dimensionalidad con Técnicas lineales y No lineales |
| 3 | Técnicas de embedding Técnicas basadas en árboles Ensemble Methods: Bagging and Random Forests |
| 4 | Técnicas de Segmentación |

| 7. MÉTODOS DE LA EVALUACIÓN | | | | |
|---|---|-------------|----------|--------|
| Descripción | Tipología | Eval. Final | Recuper. | % |
| Valoración de informes y trabajos escritos | Actividad de evaluación con soporte virtual | Sí | Sí | 60,00 |
| Examen (escrito, oral y/o práctico en el aula de computación) | Actividad de evaluación con soporte virtual | Sí | Sí | 40,00 |
| TOTAL | | | | 100,00 |
| Observaciones | | | | |
| <p>Las calificaciones mínimas se aplican a los dos bloques de evaluación, la evaluación continua y el examen final. En base al artículo 35 del Reglamento, en el caso de que un estudiante no obtuviese la calificación mínima requerida en alguna de las pruebas de evaluación, la calificación global de la asignatura será el menor valor entre 4,9 y la media ponderada de las pruebas de evaluación. Con carácter obligatorio todas las pruebas se realizarán o entregarán usando la plataforma virtual de la asignatura. Por tanto es responsabilidad del alumno, asegurarse de que pueden acceder a la plataforma virtual de la asignatura, antes del comienzo de las sesiones en las que se realicen las pruebas.</p> <p>Si la nota final del alumno fuese menor que 5 sobre 10, entonces la recuperación consistirá en la realización de cada una de las tareas en las que hubiera obtenido una calificación menor que 5 sobre 10. El procedimiento de evaluación de una actividad recuperable será equivalente al de la actividad original.</p> | | | | |
| Criterios de evaluación para estudiantes a tiempo parcial | | | | |
| Se aplicará la misma evaluación que a los estudiantes a tiempo completo, teniendo en cuenta su disponibilidad de tiempo en cuanto a los plazos de entrega de los trabajos. | | | | |

| 8. BIBLIOGRAFÍA Y MATERIALES DIDÁCTICOS |
|--|
| BÁSICA |
| <p>An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Editorial Springer. Available at http://www-bcf.usc.edu/~gareth/ISL/</p> <p>Dimension Reduction: A Guided Tour, Christopher J. C. Burges, Foundations and Trends® in Machine Learning: Vol. 2: No. 4, pp 275-365. http://dx.doi.org/10.1561/2200000002</p> <p>Machine learning: a probabilistic perspective, Kevin P. Murphy. ISBN 978-0-262-01802-9, MIT Press Cambridge, Massachusetts, London, England</p> |

Esta es la Guía Docente abreviada de la asignatura. Tienes también publicada en la Web la información más detallada de la asignatura en la Guía Docente Completa.