

Datos básicos de la materia

Carácter	<input checked="" type="checkbox"/> Obligatoria	<input type="checkbox"/> Optativa	<input type="checkbox"/> Prácticas Externas	<input type="checkbox"/> Trabajo Fin de Máster	<input type="checkbox"/> Mixta
Créditos ECTS	12 créditos ECTS				
Unidad Temporal	<input checked="" type="checkbox"/> Cuatrimestral				
ECTS Cuatrimestre 1 12	ECTS Cuatrimestre 2		ECTS Cuatrimestre 3		
Lenguas en que se imparte	<input checked="" type="checkbox"/> Castellano	<input checked="" type="checkbox"/> Inglés		<input type="checkbox"/> Otra:	

Contenidos

M02

1. Estadística descriptiva.
2. Muestreo y Monte Carlo.
3. Fundamentos de inferencia estadística: una y varias variables.
4. Contrastes paramétricos y no paramétricos para una y varias muestras.
5. Técnicas de remuestreo (bootstrap).
6. Modelos de regresión. Estimación de máxima verosimilitud.
7. Regularización. Regresión contraída (ridge regression)
8. Problemas de asociación, segmentación, clasificación, y predicción.
9. Aprendizaje basado en casos, no supervisado, supervisado, o por refuerzo.
10. Evaluación de métodos. Sobreajuste, validación cruzada (k-fold).
11. Reglas de asociación.
12. Técnicas de vecinos cercanos, distancias, núcleos y funciones de base radial.
13. Segmentación jerárquica, k-medias, SOM.
14. Árboles de clasificación y regresión.
15. Modelos lineales y aditivos generalizados.
16. Aprendizaje por conjuntos: boosting y bagging.

Observaciones

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.

Metodología:

Se comenzará por una exposición de los conceptos y métodos básicos, incluyendo ejemplos ilustrativos sencillos de distintas disciplinas, que serán analizados y discutidos en común. Los estudiantes, organizados en grupos, realizarán un análisis práctico detallado de distintos casos de estudio para problemas de asociación, segmentación, clasificación y predicción aplicando las técnicas adecuadas en cada caso. La parte práctica se realizará utilizando el software estadístico R de uso libre.

Resultados del 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.
- Entender las técnicas estadísticas de muestreo y remuestreo.
- Entender las técnicas de contraste de hipótesis y saber aplicarlas a problemas típicos.
- Entender el modelo de regresión lineal y saber aplicarlo en casos prácticos.
- Entender el método de máxima verosimilitud para estimación de parámetros.
- Conocer el concepto de regularización.
- 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 estándar disponibles en R para resolver casos prácticos de minería de datos y analizar críticamente las soluciones obtenidas.

Denominación de la materia:

M02 MÉTODOS EN CIENCIA DE DATOS/
METHODS IN DATA SCIENCE

Información sobre las asignaturas

Denominación	M02-01 Estadística para la Ciencia de Datos / Statistics for Data Science		
Carácter OBLIGATORIA	ECTS 6	Lengua impartición	castellano
Denominación	M02-02 Data Mining		
Carácter OBLIGATORIA	ECTS 6	Lengua impartición	castellano