

FICHA DE LA ASIGNATURA

Nombre de la Asignatura	Código
SEMINARIO DE MINERÍA DE DATOS Y OTRAS TÉCNICAS ESTADÍSTICAS	(no rellenar)

Módulo	(no rellenar)
Materia	(no rellenar)
Créditos para alumnos	4 ECTS
Carácter	Optativa
Prerrequisitos	
Idioma/s	Castellano
Recomendaciones	Formación previa en probabilidad y estadística multivariante. Modelos de regresión
Descriptorios	Minería de datos. Análisis multivariante. Software estadístico y análisis de datos

PROFESORES RESPONSABLES

	Profesor	e-mail
coordinador	Carlos Lamenca Martínez	clamenca@mat.ucm.es

OBJETIVOS DE LA ASIGNATURA

- Conocimientos teóricos y prácticos para la aplicación adecuada de un conjunto de técnicas estadísticas orientadas al análisis de datos.
- Capacidad de explicar distintos tipos de interrelaciones entre variables.
- Capacidad de resumir conjuntos de datos mediante un grupo reducido de variables.
- Capacidad de analizar las relaciones entre variables o entre individuos de un conjunto de datos, con el fin de formar grupos con características similares.
- Capacidad de hacer predicciones para la clasificación de nuevas observaciones.

COMPETENCIAS DE LA ASIGNATURA

Competencias generales y básicas

CG1. Aprender a aplicar los conocimientos adquiridos y a explotar su potencial para la resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares).

CG2. Elaborar adecuadamente y con originalidad argumentos motivados y proyectos de trabajo, redactar planes, así como formular hipótesis y conjeturas razonables en su área de especialización.

CG3. Integrar los conocimientos adecuados y enfrentarse a la complejidad de emitir juicios en función de criterios, de normas externas o de reflexiones personales justificadas.

CG4. Comunicar y presentar públicamente ideas, procedimientos o informes de investigación, así como asesorar a personas u organizaciones en el tratamiento estadístico.

CG5. Conocer los modelos, métodos y técnicas relevantes en distintas áreas de aplicación de la Estadística.

CG7. Habilidades de investigación e innovación: capacidad para identificar y formular hipótesis o ideas innovadoras y someterlas a prueba de objetividad y coherencia.

CG8. Capacidad de toma de decisiones, de asunción de responsabilidades en entornos complejos y de liderazgo en marcos multiculturales.

CB6. 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.

CB7. 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.

CB8. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

CB9. 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.

CB10. 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.

Competencias específicas

CE13. Manejar de forma avanzada las herramientas informáticas para el tratamiento y la representación de datos oficiales.

CE14. Estar familiarizado tanto con las técnicas estadísticas más recientes como con las tendencias de desarrollo futuro del análisis cuantitativo.

CE18. Desarrollar las capacidades necesarias para valorar, evaluar y diseñar instrumentos de medición cuantitativa.

Competencias transversales

CT1. Buscar y localizar información, utilizando para ello distintos tipos de fuentes documentales y estadísticas (bibliotecas físicas y virtuales, bases de datos, páginas webs de organismos internacionales y nacionales, públicos y privados, manuales y libros, artículos en revistas especializadas, anuarios estadísticos, etc.).

CT2. Aprender de forma continua, sabiendo auto-gestionar su estudio, en función de las necesidades de elaboración y análisis de la información.

CT4. Aplicar conocimientos generales a situaciones prácticas concretas, incluso en contextos de limitaciones o presión (temporales, de recursos, etc.).

CT5. Utilizar eficientemente las tecnologías de la información y las telecomunicaciones (Internet, aplicaciones informáticas y recursos telemáticos).

CONTENIDOS TEMÁTICOS

1. Minería de datos.
2. Regresión logística.
3. KNN.
4. Métodos de remuestreo.
5. Selección de modelos.
6. Métodos de regularización.
7. Árboles de decisión.
8. Ensamblaje de modelos.
9. Otras técnicas estadísticas.

MATERIAL BIBLIOGRÁFICO Y OTROS RECURSOS DIDÁCTICOS

Bibliografía básica

- Han J., Kamber M. y Pei J. (2012) Data Mining. Concepts and techniques. Elsevier.
- Maimon, O. y Rokach, L. (2010) Data Mining and Knowledge Discovery Handbook. Springer.
- Hastie, T., Tibshirani, R. y Friedman, J. (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
- Pérez López, C. y Santín González, D. (2007) Minería de datos: técnicas y herramientas. Thomson.
- Hernández Orallo, J., Ramírez Quintana, M.J. y Ferri Ramirez, C. (2004) Introducción a la minería de datos. Pearson Educación.

Bibliografía complementaria

- Peña, D. (2002) Análisis de datos multivariantes. McGraw-Hill.

Otros recursos

- R (<https://www.r-project.org>)

GESTIÓN DEL PROCESO DE APRENDIZAJE

ACTIVIDADES PRESENCIALES DE ALUMNOS/AS Y PROFESORES/AS	Horas previstas
Sesiones académicas teóricas	13
Sesiones académicas de problemas	13
Tutorías grupales	0
Presentaciones y debates	4
TOTAL DE HORAS PRESENCIALES	30

ACTIVIDADES DE TRABAJO AUTÓNOMO DE LOS/AS ALUMNOS/AS	Horas previstas
Preparación de seminarios, lecturas, investigaciones, etc., para exponer y entregar en las clases teóricas y prácticas	35
Preparación de trabajos finales para entregar a la conclusión del curso: trabajos de análisis y revisiones teóricas y de investigación en el campo de la comunicación (trabajos exploratorios, de recopilación de fuentes bibliográficas, estadísticas, históricas, etc.	35
Actividades formativas alternativas relacionadas con la asignatura y de interés para la formación del estudiante. Asistencia a jornadas, congresos y conferencias, etc.	0
TOTAL DE HORAS DE TRABAJO AUTÓNOMO DE LOS ALUMNOS/AS	70

EVALUACIÓN

Actividades que serán evaluadas	Calificación (% nota final)
Asistencia y participación en las clases	10%
Trabajo individual	40%
Trabajo grupal	40%
Prueba escrita final	50%