

FICHA DE ASIGNATURA

Título: Minería de datos

Descripción:

El enfoque básico del curso son los procedimientos automáticos encuadrados bajo la categoría de “minería de datos” que permiten describir y modelar conjuntos de datos complejos. Se aprenden mecanismos que detectan patrones y regularidades en dichos datos y que pueden ser usados para predecir respuestas ante nuevas situaciones, detectar los campos clave que determinan las características del problema, realizar segmentaciones de forma automática, lanzar pronósticos, detectar excepciones, proponer escenarios; en definitiva, anticipar el futuro de forma objetiva y basada en datos reales.

Carácter: Obligatoria

Créditos ECTS: 6

Contextualización:

El programa de esta asignatura proporciona al alumno conocimientos para extraer información relevante y generalmente implícita y desconocida de datos masivos y usarla para realizar pronósticos.

Modalidad: Online

Temario:

1. Proyectos de minería de datos
2. Proceso de KDD
3. Técnicas de pre-procesamiento de datos
4. Métodos de clasificación
5. Sistemas de recomendación
6. Análisis de datos con cubos y modelos de minería
7. Evaluación y selección de modelos: matriz de confusión, métricas, costes. Curvas ROC.

Competencias:

CE7: Aplicar los principales métodos de minería de datos para extraer características y contextualizar información con el fin de facilitar la posterior toma de decisiones.

CE11: Elaborar modelos de decisión estratégica basados en técnicas de clasificación y tratado de datos, haciendo uso del conocimiento específico de los requisitos y necesidades del usuario final del modelo.

CE15: Evaluar las diferentes soluciones Big Data frente a un problema y seleccionar en base a criterios de eficiencia y otros, las técnicas óptimas para cada problema, así como ser capaz de ejecutar la solución de forma adecuada e interpretar los resultados obtenidos.

CE16: Comunicar con claridad, a los grupos decisores usuarios de la información, las conclusiones obtenidas en el proceso de análisis de datos.

Actividades Formativas:

Actividad Formativa	Horas	Presencialidad
Clases expositivas	12	0%
Clases prácticas sobre laboratorio informático	12	0%
Tutorías online	10	30%
Trabajo autónomo	116	0%

Metodologías docentes:

- Lección magistral
- Estudio de casos
- Resolución de problemas
- Simulaciones
- Laboratorio informático virtual
- Seguimiento

Sistema de Evaluación:

Sistemas de evaluación	Ponderación mínima	Ponderación máxima
Evaluación de portafolio. Informe sobre resolución de problemas o estudio de casos	20%	30%
Evaluación del portafolio. Informe sobre Diseño de Proyectos	10%	20%
Participación en grupos de debate	10%	10%
Evaluación de la prueba	40%	60%

Normativa específica:

No procede

Bibliografía:

- Abbas, H.A., Sarker, R.A., Newton, C.S. (2002), *Data mining. A heuristic approach*. Idea Group Publishing, Hershey.
- Efron, B., Hastie, T. (2016), *Computer age statistical inference: algorithms, evidence, and data science*. Cambridge University Press, Cambridge.
- Hastie, T., Tibshirani, R., Friedman, L. (2009), *The elements of statistical learning*. Springer, Berlín.

- Hernández, J., Ramírez, M.J., Ferri, C. (2004), *Introducción a la Minería de Datos*. Pearson-Prentice-Hall, Madrid.
- James, G., Witten, D., Hastie, T. (2013), *Introduction to Statistical Learning with Applications in R*. Springer, Berlín.
- Khun, M., Johnson, K. (2013), *Applied Predictive Modeling*. Springer, Berlín.
- Lara, J.A. (2014), *Minería de datos*. Centro de Estudios Financieros, Madrid.
- Pérez, M. (2014), *Minería de datos a través de ejemplos*. RC Libros
- Sarker, R.A., Abbas, H.A., Newton, C.S. (2002), *Heuristics & optimization for knowledge Discovery*. Idea Group Publishing, Hershey.
- Siegel, E. (2013), *Analítica predictiva. Predecir el futuro utilizando big data*. Anaya, Madrid.