

Presentación de la asignatura Tratamiento Inteligente de Datos



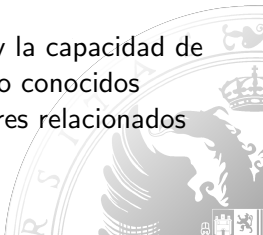
Maria-Amparo Vila
vila@decsai.ugr.es

Grupo de Investigación en Bases de
Datos y Sistemas de Información
Inteligentes <https://idbis.ugr.es/>
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Universidad de Granada

Introducción a la asignatura

Competencias y destrezas básicas

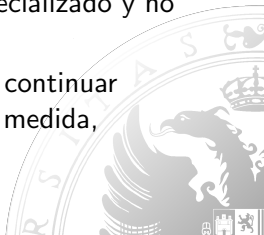
- G4.- Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática.
- CB6.- Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- CB7.- Saber aplicar los conocimientos adquiridos y la capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios o multidisciplinarios relacionados con diferentes áreas de estudio o trabajo.



Introducción a la asignatura

Competencias y destrezas básicas

- CB8.- Capacidad de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
- CB9.- Capacidad para comunicar conclusiones, y los conocimientos y razones últimas que las sustenten, a público especializado y no especializado, de forma clara y sin ambigüedades.
- CB10.- Capacidad y habilidad de aprendizaje para continuar estudiando de un modo que habrá de ser, en gran medida, autodirigido o autónomo.



Introducción a la asignatura

Competencias y destrezas Específicas

- TI5.- Capacidad para analizar las necesidades de información que se plantean en un entorno y llevar a cabo en todas sus etapas el proceso de construcción de un sistema de información.
- TI9.- Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento.



Introducción a la asignatura

Competencias y destrezas Transversales

- T1.- Capacidad de Análisis y síntesis. Capacidad para encontrar, analizar, criticar, relacionar, estructurar y sintetizar información proveniente de diversas fuentes, así como integrar ideas y conocimientos
- T2.- Capacidad de organización y planificación, así como capacidad de gestión de la información.
- T3.- Capacidad para el uso de las TIC en el ámbito académico y profesional.
- T6.- Capacidad para innovar y generar nuevas ideas



Introducción a la asignatura

Objetivos genéricos

- Aplicar los métodos básicos para la limpieza de datos necesaria, previa a cualquier análisis estadístico o destinado al aprendizaje automático.
- Analizar y categorizar adecuadamente algoritmos y modelos de aprendizaje existentes. Además, deberá adquirir el bagaje suficiente para proponer métodos alternativos adecuados a situaciones específicas no descritas anteriormente.
- Conocer las nuevas tendencias en Minería de Datos
- Conocer los desarrollos más importantes de Minería de Datos Difusa.

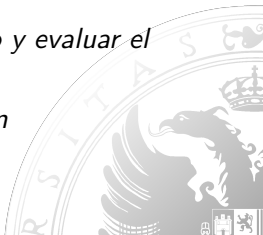


Introducción a la asignatura

Objetivo Global

Que los alumnos sean capaces de:

- *Detectar los problemas de Análisis Inteligente de Datos en su ámbito profesional*
- *Seleccionar y preparar conjuntos de datos para resolver dichos problemas*
- *Reconocer el modelo a aplicar*
- *Aplicar las técnicas más apropiadas para resolverlo y evaluar el resultado*
- *Interpretar (en su caso) adecuadamente la solución*



Introducción a la asignatura

Asignaturas con las que se relaciona

Asignaturas de formación básica del grado

Estadística

Asignaturas de especialidad del grado:

Aprendizaje Automático

Inteligencia de Negocio

Asignaturas del master con las que se relaciona

Se apoya en:

Inteligencia Computacional

Sirve para:

Sistemas Inteligentes para la gestión en la empresa

Gestión de la Información en la Web



Introducción a la asignatura

Programa de la asignatura

Tema 1 *Introducción al Tratamiento Inteligentes de Datos (IDA)*

Motivación histórica. IDA y KDD. IDA y Minería de Datos (DM). DM y Estadística. DM y Aprendizaje. Problemas más importantes en DM.

Tema 2 *Los datos en IDA.*

Tipos de datos. Limpieza y visualización. Transformación de coordenadas: componentes principales. Cambios de escala.



Introducción a la asignatura

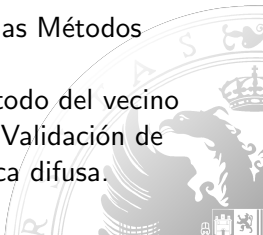
Programa de la asignatura

Tema 3 *Problemas de agrupamiento (Clustering)*

Introducción, tipos de datos y distancias. Clustering jerárquico. Métodos particionales: kmedias, kmedoides, fuzzyKmedias. Métodos de análisis de densidad: DBSCAN. Problemas de escala en CLustering. Validación de agrupamientos

Tema 4 *Problemas de clasificación.*

Introducción. Árboles de decisión y reglas Métodos Bayesianos: Naïve Bayes y Análisis Discriminante. Métodos predictivos: método del vecino más cercano. redes neuronales y SVM. Validación de resultados. Clasificación basada en lógica difusa.



Introducción a la asignatura

Programa de la asignatura

Tema 5 *Problemas de asociación entre atributos. Regresión*
Introducción: ideas básicas. La regresión Lineal. La regresión multivariante. La regresión logística.

Tema 6 *Problemas de asociación entre atributos. Reglas de asociación*

Introducción. Reglas de asociación: definición, mecanismos de cálculo. Medidas de bondad de RA. Introducción a las reglas de asociación difusas. Relaciones funcionales entre atributos: análisis de regresión.

Tema 7 *Problemas de Minería de textos*

Introducción de la Minería de Textos (TM). Preprocesamiento de datos textuales. Análisis de semántica latente. Clustering y asociación en TM.



Introducción a la asignatura

Programa de la asignatura

Tema 8 *Predicción en el tiempo*

Minería de series temporales. Minería de secuencias.



Introducción a la asignatura

Programa de la asignatura: bibliografía básica

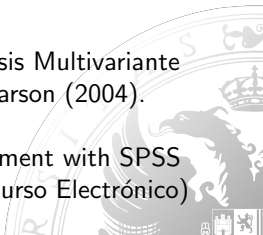
- **Teoría**

- Pang-Ning Tan, Michael Steinbach , Vipin Kumar Introduction to Data Mining Addison Wesley (2006) (Biblioteca)
- Michael R. Berthold, Christian Borgelt, Frank Höppner, Frank Klawonn Guide to Intelligent Data Analysis Springer Verlag (2010) (Recurso electrónico)
- M.J.Zaki, W. Meira Data Mining and Analysis. Fundamental Concepts and Algorithms Cambridge Univ. Press. (2014) Recurso Electrónico

- **Prácticas**

- Para SPSS

- Cesar Perez Lopez Técnicas de Análisis Multivariante de Datos. Aplicaciones con SPSS Pearson (2004). (Biblioteca)
- J.P. Verma Data Analysis in Management with SPSS software Springer Verlag (2013) (Recurso Electrónico)



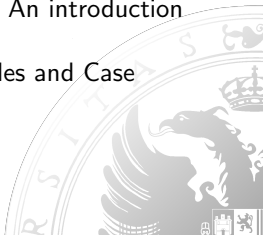
Introducción a la asignatura

Programa de la asignatura: bibliografía básica

- **Prácticas**

- Para **Knime**
 - G.Bákos KNIME Essential PACK PUB. 2015 (Recurso Electrónico Biblioteca)
 - Distintas guías de inicio: Quick Start Knime, Primeros Pasos Knime, Capítulos del libro "Knime beginner's luck" (Recurso electrónico)

- Para **R (Rstudio)**
 - W.N. Venables and D.M. Smith An introduction to R (recurso electrónico)
 - Y. Zhao R and Data Mining: Examples and Case Studies (Recurso electrónico)



Introducción a la asignatura

Mecanismos de evaluación

- Asistencia y participación (10% (aprox))
- Examen tipo test de conceptos básicos. (35% (aprox))
- Trabajo practico individual. (55%(aprox))

*El trabajo consistirá en el **análisis completo** de un conjunto de datos seleccionado por el alumno y en la **exposición oral** de los resultado*

Temporización: (tentativa)

- Hasta el 10 de Noviembre presentación de propuestas
- Desarrollo tutelado del 11 de Noviembre al 16 de Enero
- Presentación de los trabajos a partir del 8 de Enero
- Examen de conceptos el día 22 de Enero

En los puntos 2 y 3 de la evaluación será necesario obtener más de 3,5 para hacer la media

