GUÍA DOCENTE

DENOMINACIÓN DE LA ASIGNATURA

Denominación: **ANÁLISIS AUTOMÁTICO DE DATOS PARA LAS CIENCIAS BIOMÉDICAS, MEDIOAMBIENTALES,**
Código: 138014
Plan de estudios: **TRANSVERSALES MÁSTERES UNIVERSITARIOS** Curso: 1
Créditos ECTS: 4.0 Horas de trabajo presencial: 30
Porcentaje de presencialidad: 30% Horas de trabajo no presencial: 70
Plataforma virtual: http://www.uco.es/moodle

DATOS DEL PROFESORADO

Nombre: HERVAS MARTINEZ, CESAR (Coordinador)
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: CAMPUS DE RABANALES, EDIFICIO ALBERT EINSTEIN (C2), 3a PLANTA
E-Mail: chervas@uco.es Teléfono: 957218349
URL web: http://www.uco.es/ayrna

Nombre: FERNÁNDEZ CABALLERO, JUAN CARLOS
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: CAMPUS DE RABANALES, EDIFICIO MARIE CURIE (C3), ANEXO
E-Mail: jfcaballero@uco.es Teléfono: 957218153
URL web: http://www.uco.es/ayrna

Nombre: GUTIÉRREZ PEÑA, PEDRO ANTONIO
Departamento: INFORMÁTICA Y ANÁLISIS NUMÉRICO
Área: CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
Ubicación del despacho: CAMPUS DE RABANALES, EDIFICIO ALBERT EINSTEIN (C2), 3a PLANTA
E-Mail: pagutierrez@uco.es Teléfono: 957218153
URL web: http://www.uco.es/~i02gupep

REQUISITOS Y RECOMENDACIONES

Requisitos previos establecidos en el plan de estudios

Ninguno.

Recomendaciones

Se recomienda tener conocimientos básicos de Estadística.

COMPETENCIAS

CTEC5	Capacidad para adquirir, obtener, formalizar y representar el conocimiento humano en una forma computable para la resolución de problemas mediante un sistema informático en cualquier ámbito de aplicación
CTEC7	Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y utilizar aplicaciones y sistemas que las implementen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de pequeños, medianos y grandes volúmenes de datos
CU2	Conocer y perfeccionar el nivel de usuario en el ámbito de las Tecnologías de la Información y las Comunicaciones

GUÍA DOCENTE

OBJETIVOS

La asignatura pretende acercar al alumnado al aprendizaje automático y su aplicabilidad en distintas áreas. Para ello, se plantean los siguientes objetivos:
- Construcción de una base de datos real u obtenida del repositorio UCI (Universidad de California en Irvine).
- Preprocesado de la base de datos (técnicas de minería de datos como selección de características, detección de outliers, etc).
- Comprensión de las técnicas básicas de aprendizaje automático para regresión, clasificación y agrupamiento (clustering).
- Utilización práctica de las técnicas básicas en problemas reales.
- Análisis de las metodologías más adecuadas para cada técnica y de los resultados obtenidos mediante las mismas.
- Extracción de conclusiones y discusión de resultados y de las posibles mejoras de los experimentos.

CONTENIDOS

1. Contenidos teóricos

Tema 1. Presentación / Introducción al aprendizaje automático y a WEKA.
Tema 2. Ejemplos de aplicaciones en problemas reales.
Tema 3. Preprocesamiento: tipos de variables y conversión de datos, selección de características, limpieza de datos.
Tema 4. Aprendizaje no supervisado: Agrupamiento. Ejemplos
Tema 5. Aprendizaje Supervisado. Clasificación: medidas de evaluación. Ejemplos
Tema 6. Clasificación: regresión logística. Ejemplos
Tema 7. Clasificación: redes neuronales. Ejemplos

2. Contenidos prácticos

Utilización de las aplicaciones software Weka y SPSS para aplicar las metodologías abordadas en la parte teórica. Aplicación a bases de datos reales.

METODOLOGÍA

Aclaraciones

Lecciones magistrales: En estas sesiones se presentarán los conceptos de cada tema a la par que se desarrollarán ejemplos que ayuden al alumno a comprender y clarificar los conceptos. Durante las clases teóricas el profesor propondrá y realizará algunos ejercicios prácticos, que exploten la teoría abordada.

Laboratorio: Las clases prácticas se desarrollarán mediante presentaciones que versarán sobre las metodologías descritas en los objetivos. Durante las clases prácticas, el profesor controlará que los alumnos han comprendido las metodologías abordadas. El objetivo fundamental de estas clases consiste en la realización de un trabajo práctico.

Actividades presenciales

Actividad	Total
Laboratorio	15

GUÍA DOCENTE

Actividad	Total
Lección magistral	15
Total horas:	***30***

Actividades no presenciales

Actividad	Total
Búsqueda de información	10
Consultas bibliográficas	20
Trabajo práctico	40
Total horas:	***70***

MATERIAL DE TRABAJO PARA EL ALUMNO

Enlaces web interesantes - *http://www.uco.es/moodle*
Manual de la asignatura - *http://www.uco.es/moodle*
Presentaciones PowerPoint - *http://www.uco.es/moodle*

Aclaraciones

Para facilitar el estudio personal del alumno, la asimilación de los contenidos y el desarrollo de las competencias y habilidades, el alumno tendrá a su disposición el material de trabajo indicado.

EVALUACIÓN

Instrumentos	Porcentaje
Asistencia (lista de control)	30%
Pruebas orales	20%

GUÍA DOCENTE

Periodo de validez de las calificaciones parciales:

Las calificaciones parciales obtenidas se guardarán durante un curso académico.

Aclaraciones:

El **trabajo práctico** se valorará atendiendo a tres criterios: presentación (portada, paginación, bibliografía, etc.), dificultad (utilización de una o varias metodologías de aprendizaje automático) e interés (novedad de la base de datos utilizada, interpretación de los modelos obtenidos en dicha base de datos, etc.). El trabajo práctico tendrá que tener una calificación mínima de **5 puntos sobre 10** para hacer media.

Las **pruebas orales** consistirán en la defensa del trabajo y resolución de las preguntas planteadas por el profesor. La prueba oral tendrá que tener una calificación mínima de **5 puntos sobre 10** para hacer media.

La lista de control sirve para evaluar la asistencia a clase del alumno. La calificación en este apartado será proporcional al número de días de asistencia a clase.

BIBLIOGRAFIA

1. Bibliografía básica

- C. M. Bishop, Pattern recognition and machine learning, Springer, 2006.
- Ian H. Witten, Eiben Frank. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann, 1999.
- J. Hernández-Orallo, M.J. Ramírez, C. Ferri. Introducción a la Minería de Datos. Pearson Educación. 2004.

2. Bibliografía complementaria

- Tom M. Mitchell. Machine Learning. McGraw-Hill, 1997.
- Jiawei Han and Micheline Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2000.
- Basilio Sierra Araujo. Aprendizaje Automático: conceptos básicos y avanzados. Aspectos prácticos utilizando el software WEKA. Pearson Education. Prentice Hall. 2006
- Alpaydin, E. 2010 Introduction to Machine Learning, 2Ed. The MIT Press.
- R. Duda, P.E. Hart, D.G. Stork. Pattern Classification. Wiley. 2001.
- L. Kuncheva. Combining Pattern Classifiers. Wiley. 2004. S. Sharma. Applied Multivariate Techniques. Wiley. 1996.
- Hastie, T. and Tibshirani, R. and Friedman. 2001 The elements of statistical learning: data mining, inference, and prediction. Springer.
- A. Webb. Statistical Pattern Recognition. Wiley. 2002.
- Software WEKA: http:// www.cs.waikato.ac.nz/ml/weka.

Las estrategias metodológicas y el sistema de evaluación contempladas en esta Guía Docente serán adaptadas de acuerdo a las necesidades presentadas por estudiantes con discapacidad y necesidades educativas especiales en los casos que se requieran.