



UNIVERSIDAD DE CASTILLA - LA MANCHA

GUÍA DOCENTE

1. DATOS GENERALES

Asignatura: MINERÍA DE DATOS

Tipología: OPTATIVA

Grado: 406 - GRADO EN INGENIERÍA INFORMÁTICA (AB)_20

Centro: 604 - E.S. DE INGENIERIA INFORMÁTICA ALBACETE

Curso: 4

Lengua principal de impartición: Español

Uso docente de otras lenguas: Las clases de teoría y prácticas se darán en español, no obstante, la gran mayoría del material (diapositivas, enunciados de prácticas y problemas, etc.) estarán en inglés al tratarse de una asignatura English-Friendly.

Página web:

Código: 42348

Créditos ECTS: 6

Curso académico: 2023-24

Grupo(s): 15

Duración: Primer cuatrimestre

Segunda lengua: Inglés

English Friendly: S

Bilingüe: N

Profesor: JOSE ANTONIO GAMEZ MARTIN - Grupo(s): 15				
Edificio/Despacho	Departamento	Teléfono	Correo electrónico	Horario de tutoría
ESIV/1.C.13	SISTEMAS INFORMÁTICOS	2473	jose.gamez@uclm.es	lunes 10:15-13:15 y martes 17:00-20:00

2. REQUISITOS PREVIOS

Para cursar esta materia es aconsejable haber cursado los módulos de Formación Básica (Módulo I) y el módulo Común a la Rama de Informática (Módulo II).

Esta asignatura se apoya en las competencias y los conocimientos adquiridos en las asignaturas:

- Lógica
- Estadística
- Diseño de Algoritmos
- Sistemas Inteligentes
- Sistemas Basados en el Conocimiento

3. JUSTIFICACIÓN EN EL PLAN DE ESTUDIOS, RELACIÓN CON OTRAS ASIGNATURAS Y CON LA PROFESIÓN

La asignatura se ubica en la intensificación de Computación, en el ámbito de las asignaturas de la Inteligencia Artificial. La Minería de datos, y el aprendizaje automático en general, están ligados al campo de la estadística y de la algorítmica, y abordan las técnicas para la extracción de conocimiento implícito en conjuntos de datos. En los últimos años, estas disciplinas están ganando importancia debido al incremento en la producción de datos -propiciado por fenómenos como por ejemplo el auge de internet o las redes sociales - o el desarrollo de nuevas técnicas para la obtención de información genética. Desde el punto de vista profesional, cada vez existe una mayor demanda de analistas de datos en ámbitos tan diversos como el marketing, el análisis de mercados, la seguridad, o la biología.

4. COMPETENCIAS DE LA TITULACIÓN QUE LA ASIGNATURA CONTRIBUYE A ALCANZAR

Competencias propias de la asignatura

Código	Descripción
CM05	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, particularmente los relacionados con aspectos de computación, percepción y actuación en ambientes o entornos inteligentes.
CM07	Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.
INS05	Capacidad para argumentar y justificar lógicamente las decisiones tomadas y las opiniones.
UCLM03	Correcta comunicación oral y escrita.

5. OBJETIVOS O RESULTADOS DE APRENDIZAJE ESPERADOS

Resultados de aprendizaje propios de la asignatura

Descripción

Desarrollar e implementar un sistema de recuperación de la información de tamaño pequeño a mediano.

Conocer y desarrollar técnicas de aprendizaje computacional, tanto supervisado como no supervisado, y diseñar e implementar aplicaciones y sistemas que las utilicen.

Describir y aplicar las distintas fases del proceso de descubrimiento de extracción del conocimiento a partir de grandes volúmenes de datos.

Resultados adicionales

Obtener resultados concluyentes fruto del proceso de extracción del conocimiento y ser capaz de presentarlos y justificarlos

6. TEMARIO

Tema 1: Introducción a la minería de datos.

Tema 2: Proceso de descubrimiento de conocimiento a partir de datos.

Tema 3: Validación y evaluación de modelos.

Tema 4: Métodos basados en analogía (kNN)

Tema 5: Predicción numérica: regresión.

Tema 6: Árboles: clasificación y regresión.

Tema 7: Clasificadores probabilísticos.

Tema 8: Reducción de la dimensionalidad.

Tema 9: Multclasificadores (ensembles).

Tema 10: Redes neuronales.

Tema 11: Clustering

Tema 12: Reglas de asociación.

7. ACTIVIDADES O BLOQUES DE ACTIVIDAD Y METODOLOGÍA							
Actividad formativa	Metodología	Competencias relacionadas (para títulos anteriores a RD 822/2021)	ECTS	Horas	Ev	Ob	Descripción
Enseñanza presencial (Teoría) [PRESENCIAL]	Método expositivo/Lección magistral	CM05 CM07 INS05	1.26	31.5	N	-	Se usará para introducir los principales conceptos de cada tema.
Prueba final [PRESENCIAL]	Pruebas de evaluación	CM05 CM07 INS05 UCLM03	0.1	2.5	S	S	Corresponde al examen oficial de la asignatura. De carácter individual. Es la misma prueba para evaluación continua y no continua. En convocatoria extraordinaria se recupera con el examen convocado a tal efecto por jefatura de estudios.
Talleres o seminarios [PRESENCIAL]	Método expositivo/Lección magistral	CM05 CM07 INS05	0.08	2	N	-	Se realizará un seminario en laboratorio de las herramientas/APIs usadas para la implementación de las prácticas e ilustración de los casos de estudio (Python, Weka, etc). Corresponde a la práctica 0.
Enseñanza presencial (Prácticas) [PRESENCIAL]	Método expositivo/Lección magistral	CM05 CM07 INS05	0.06	1.5	N	-	Se usará la primera media hora de la primera sesión correspondiente a cada una de las tres prácticas obligatorias para describir el trabajo a realizar, el material disponible y el entregable a presentar.
Prácticas en aulas de ordenadores [PRESENCIAL]	Trabajo dirigido o tutorizado	CM05 CM07 INS05	0.66	16.5	N	-	Trabajo del alumno con supervisión del profesor en las prácticas de la asignatura. Será la base para realizar las prácticas a entregar.
Resolución de problemas o casos [PRESENCIAL]	Resolución de ejercicios y problemas	CM05 CM07 INS05	0.24	6	N	-	Resolución en clase de problemas y casos de estudio relacionados con los distintos temas estudiados.
Estudio o preparación de pruebas [AUTÓNOMA]	Trabajo autónomo	CM05 CM07 INS05	1.56	39	N	-	Se estima que el alumno dedicará aproximadamente 1.3 horas de estudio autónomo por cada hora de lección magistral recibida. Esto incluye la búsqueda y lectura de material adicional.
Otra actividad no presencial [AUTÓNOMA]	Prácticas	CM05 CM07 INS05	0.84	21	N	-	Horas adicionales a las presenciales para completar las prácticas/proyectos de programación.
Elaboración de memorias de Prácticas [AUTÓNOMA]	Trabajo autónomo	CM05 CM07 INS05 UCLM03	0.42	10.5	S	S	Los alumnos han de plasmar en una memoria el trabajo realizado en la práctica, presentar de forma coherente los resultados y obtener las conclusiones de acuerdo a los objetivos inicialmente fijados. Esta memoria podrá ser solicitada como documento independiente o directamente incrustada en los distintos apartados de las libretas a entregar con el código de cada práctica (3). Las prácticas se diseñan para ser realizadas y defendidas por parejas. No obstante en situaciones particulares y pactadas de antemano con el profesor, podrán realizarse de forma individual. En caso de no superarse las prácticas en evaluación continua, se podrán recuperar en la entrega correspondiente a la convocatoria ordinaria, si bien, el enunciado podrá ser ligeramente diferente. La convocatoria extraordinaria y de finalización se rige por las

Elaboración de informes o trabajos [AUTÓNOMA]	Aprendizaje basado en problemas (ABP)	CM05 CM07 INS05 UCLM03	0.78	19.5	S	N	condiciones de la evaluación no continua. Se facilitarán casos de estudio o ejercicios seleccionados para que el alumno trabaje sobre ellos. La entrega y comentarios sobre los mismos en los foros, de carácter voluntario, será evaluada como participación en clase. Esta actividad es de carácter individual. Esta actividad no es obligatoria y no se puede recuperar. No obstante su porcentaje será sumado al del examen en modalidad no continua y en convocatoria ordinaria, por lo que el alumno se examinará sobre el 100% (entre teoría y prácticas) y podrá obtener la puntuación máxima.
Total:			6	150			
Créditos totales de trabajo presencial: 2.4			Horas totales de trabajo presencial: 60				
Créditos totales de trabajo autónomo: 3.6			Horas totales de trabajo autónomo: 90				

Ev: Actividad formativa evaluable

Ob: Actividad formativa de superación obligatoria (Será imprescindible su superación tanto en evaluación continua como no continua)

8. CRITERIOS DE EVALUACIÓN Y VALORACIONES			
Sistema de evaluación	Evaluación continua	Evaluación no continua*	Descripción
Valoración de la participación con aprovechamiento en clase	10.00%	0.00%	(ESC) Este 10% de la nota corresponde a la participación en clase y en los foros de campus virtual. El objetivo es resolver, compartir y discutir en clase y/o en los foros de la asignatura, los problemas (ejercicios o casos) que se propongan para realizar. Esta es de carácter individual y no es obligatoria ni tiene mínimo.
Elaboración de memorias de prácticas	15.00%	15.00%	(INF) Es obligatoria la entrega de las tres memorias de prácticas. Se abrirá plazo de entrega específico para la evaluación no continua. Se exige un mínimo de 4 sobre 10 en la nota promediada de las tres prácticas.
Realización de prácticas en laboratorio	15.00%	15.00%	(LAB) Se supervisará la eficiencia y eficacia del código generado para las tres prácticas planteadas. Entrega por grupo de prácticas. Se abrirá plazo de entrega específico para la evaluación no continua. Se exige un mínimo de 4 sobre 10 en la nota promediada de las tres prácticas.
Otro sistema de evaluación	15.00%	15.00%	(PRES) Cada grupo de prácticas deberá presentar su práctica al profesor y responder a las preguntas realizadas. Entrevista por grupo de prácticas, con preguntas individuales, por lo que la nota puede ser diferente para los dos miembros de la pareja. Se abrirá plazo de entrega específico para la evaluación no continua. Se exige un mínimo de 4 sobre 10 en la nota promediada de las tres prácticas.
Prueba final	45.00%	55.00%	(ESC) En la convocatoria ordinaria en la modalidad de evaluación continua esta prueba tendrá un valor del 45% de la nota final de la asignatura. En la convocatoria ordinaria en la modalidad no continua su peso será del 55%. Entrega individual. El temario es el mismo en ambas modalidades (continua y no-continua).
Total:	100.00%	100.00%	

* En **Evaluación no continua** se deben definir los porcentajes de evaluación según lo dispuesto en el art. 4 del Reglamento de Evaluación del Estudiante de la UCLM, que establece que debe facilitarse a los estudiantes que no puedan asistir regularmente a las actividades formativas presenciales la superación de la asignatura, teniendo derecho (art. 12.2) a ser calificado globalmente, en 2 convocatorias anuales por asignatura, una ordinaria y otra extraordinaria (evaluándose el 100% de las competencias).

Criterios de evaluación de la convocatoria ordinaria:

Evaluación continua:

- La participación en clase y mediante los foros asume aportar soluciones novedosas y discusiones críticas sobre las ya presentadas. En ocasiones se entregarán ejercicios realizados durante la clase. Esta actividad no es obligatoria, por lo que puede superarse la asignatura sin realizarla.
- Las prácticas deben ser entregadas y defendidas en las fechas asignadas a cada práctica. Se requiere entregar las tres prácticas y obtener un mínimo de 4 sobre 10 en la nota promediada.
- La prueba de teoría se realizará en la fecha de la convocatoria ordinaria asignada por el centro. Será necesario superarla con nota ≥ 4 .
- Si se superan los mínimos en teoría y prácticas, la calificación de la asignatura es:

$$0.45 * \text{teoría} + 0.45 * \text{prácticas} + 0.1 * \text{participación}$$

En caso contrario, es decir, alguno de los mínimos no se supera, será:

mínimo(4.0, nota-teoría) si el examen de teoría se realiza, o "No Presentado" si no se hace el examen de teoría.

Originalidad: La entrega de cualquier ejercicio (examen, memoria de prácticas, código, problemas, etc.) supone implícitamente la declaración de originalidad por parte de los autores, por lo que en caso de detección de plagio, copia, etc. se iniciarán las medidas disciplinarias oportunas.

NOTA: ¿Por defecto, el estudiante será evaluado por evaluación continua. Si desea cambiarse a evaluación no continua, debe indicarlo a través del siguiente enlace <https://www.esiiaab.uclm.es/alumnos/evaluacion.php> antes de la finalización del periodo lectivo del cuatrimestre y siempre y cuando no se haya evaluado el 50% o más de la asignatura por evaluación continua.¿

Evaluación no continua:

El alumno puede examinarse del 100% de la nota, evaluado mediante el examen de teoría y las prácticas.

Para las prácticas se abrirá un plazo de entrega específico para esta convocatoria y posteriormente se anunciarán los horarios para su defensa. El enunciado de las prácticas puede ser ligeramente diferente al de la evaluación continua.

Las reglas de evaluación son las mismas de la evaluación continua excepto los porcentajes asignados a cada parte, que ahora son:

$$0.55 * \text{teoría} + 0.45 * \text{prácticas}$$

Particularidades de la convocatoria extraordinaria:

Se seguirá el mismo esquema que para la convocatoria ordinaria en evaluación no continua. Se abrirá plazo para la entrega y defensa de las prácticas, cuyo enunciado será el mismo de la convocatoria ordinaria en evaluación no continua.

- Se considerará la nota de las actividades superadas en convocatoria ordinaria (examen de teoría y prácticas individualmente).

Particularidades de la convocatoria especial de finalización:

Se seguirá el mismo esquema que para la convocatoria ordinaria en evaluación no continua.

- Las prácticas a evaluar corresponden a las del curso anterior. Se abrirá un plazo específico para la entrega y defensa de las mismas.

9. SECUENCIA DE TRABAJO, CALENDARIO, HITOS IMPORTANTES E INVERSIÓN TEMPORAL

No asignables a temas

Horas	Suma horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	31.5
Prueba final [PRESENCIAL][Pruebas de evaluación]	2.5
Talleres o seminarios [PRESENCIAL][Método expositivo/Lección magistral]	2
Enseñanza presencial (Prácticas) [PRESENCIAL][Método expositivo/Lección magistral]	1.5
Prácticas en aulas de ordenadores [PRESENCIAL][Trabajo dirigido o tutorizado]	16.5
Resolución de problemas o casos [PRESENCIAL][Resolución de ejercicios y problemas]	6
Estudio o preparación de pruebas [AUTÓNOMA][Trabajo autónomo]	39
Otra actividad no presencial [AUTÓNOMA][Prácticas]	21
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo autónomo]	10.5
Elaboración de informes o trabajos [AUTÓNOMA][Aprendizaje basado en problemas (ABP)]	19.5

Comentarios generales sobre la planificación: La planificación es ORIENTATIVA, pudiendo variar a lo largo del periodo lectivo en función de las necesidades docentes, festividades, o por cualquier otra causa imprevista. La planificación semanal de la asignatura podrá encontrarse de forma detallada y actualizada en la plataforma Campus Virtual (Moodle). Las actividades presenciales se organizan en tres clases de 1.5 horas por semana. Las clases concretas a usar para cubrir los 6 créditos (60 horas presenciales) se anunciarán oportunamente en CampusVirtual.

Actividad global

Actividades formativas	Suma horas
Prueba final [PRESENCIAL][Pruebas de evaluación]	2.5
Talleres o seminarios [PRESENCIAL][Método expositivo/Lección magistral]	2
Enseñanza presencial (Prácticas) [PRESENCIAL][Método expositivo/Lección magistral]	1.5
Prácticas en aulas de ordenadores [PRESENCIAL][Trabajo dirigido o tutorizado]	16.5
Resolución de problemas o casos [PRESENCIAL][Resolución de ejercicios y problemas]	6
Estudio o preparación de pruebas [AUTÓNOMA][Trabajo autónomo]	39
Otra actividad no presencial [AUTÓNOMA][Prácticas]	21
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo autónomo]	10.5
Elaboración de informes o trabajos [AUTÓNOMA][Aprendizaje basado en problemas (ABP)]	19.5
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	31.5
Total horas:	150

10. BIBLIOGRAFÍA, RECURSOS

Autor/es	Título/Enlace Web	Editorial	Población	ISBN	Año	Descripción
Xindong Wu, Vipin Kumar	The Top Ten Algorithms in Data Mining Manuales de Python.	Chapman and Hall/CRC		9781420089646	2009	
García, Salvador, Luengo, Julián, Herrera, Francisco	Data Preprocessing in Data Mining	Springer		978-3-319-10246-7	2015	
José Hernández Orallo, M.José Ramírez Quintana, Cèsar Ferri Ramírez	INTRODUCCIÓN A LA MINERÍA DE DATOS	Pearson		84 205 4091 9	2004	
Pang-Ning Tan, Michael Steinbach, and Vipin Kumar	Introduction to Data Mining	Addison-Wesley Longman Publishing Co		0321321367	2005	
Witten, Frank & Hall	Data Mining: Practical Machine Learning Tools and Techniques	Morgan & Kauffmann		978-0-12-374856-0	2011	
Joel Grus	Data Science from Scratch: First Principles with Python	O'Reilly UK Ltd		978-1492041139	2019	
Aurélien Géron	Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition https://learning.oreilly.com/library/view/hands-on-machine-learning/9781492032632/	O'Reilly Media, Inc.		9781492032649	2019	