



UNIVERSIDAD DE CASTILLA - LA MANCHA

GUÍA DOCENTE

1. DATOS GENERALES

Asignatura: MINERÍA DE DATOS

Tipología: OBLIGATORIA

Grado: 347 - GRADO EN INGENIERÍA INFORMÁTICA (CR)

Centro: 108 - ESCUELA SUPERIOR DE INFORMÁTICA C. REAL

Curso: 4

Lengua principal de impartición: Español

Uso docente de otras lenguas:

Página web: Espacio virtual de la asignatura en <https://campusvirtual.uclm.es>

Código: 42348

Créditos ECTS: 6

Curso académico: 2023-24

Grupo(s): 20

Duración: Primer cuatrimestre

Segunda lengua:

English Friendly: S

Bilingüe: N

Profesor: FERNANDO GUALO CEJUDO - Grupo(s): 20				
Edificio/Despacho	Departamento	Teléfono	Correo electrónico	Horario de tutoría
	TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN		Fernando.Gualo@uclm.es	
Profesor: ARTURO PERALTA MARTIN-PALOMINO - Grupo(s): 20				
Edificio/Despacho	Departamento	Teléfono	Correo electrónico	Horario de tutoría
FERMIN CABALLERO	TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN	926295300	Arturo.Peralta@uclm.es	

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.
CM7	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.
INS01	Capacidad de análisis, síntesis y evaluación.
INS04	Capacidad de resolución de problemas aplicando técnicas de ingeniería.
INS05	Capacidad para argumentar y justificar lógicamente las decisiones tomadas y las opiniones.
PER02	Capacidad de trabajo en equipo interdisciplinar.
PER04	Capacidad de relación interpersonal.
PER05	Reconocimiento a la diversidad, la igualdad y la multiculturalidad.
SIS01	Razonamiento crítico.
SIS03	Aprendizaje autónomo.
SIS09	Tener motivación por la calidad.
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.

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

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

6. TEMARIO

Tema 1: Introducción

Tema 1.1 Inteligencia Artificial, KDD y Data Mining.

Tema 1.2 Análisis Inteligente de Datos en entornos Big Data

Tema 2: Análisis de Datos Exploratorio

Tema 3: Tareas de Minería de Datos

Tema 3.1 Agrupamiento

Tema 3.2 Reducción de la Dimensionalidad

Tema 3.3 Extracción de Reglas de Asociación

Tema 3.4 Detección de Anomalías

Tema 3.5 Clasificación

Tema 3.6 Regresión

Tema 4: Aplicaciones de la Minería de Datos

Tema 5: Casos de Estudio

COMENTARIOS ADICIONALES SOBRE EL TEMARIO

PRÁCTICAS

En las prácticas, a lo largo del curso se desarrollará un proceso KDD completo. El alumno propondrá el dominio vinculado a sus intereses de trabajo y/o investigación o algún tema propuesto por los profesores.

1. Elección de la BD
2. Selección, tarjeta de datos.
3. Preproceso.
4. Transformación.
5. Minería de Datos.
6. Uso de los patrones descubiertos en una aplicación.

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 CM7	0.6	15	N	-	Exposición del temario por parte del profesor (MAG)
Tutorías individuales [PRESENCIAL]		CM05 CM7 INS05 SIS01 SIS09 UCLM03	0.18	4.5	N	-	Tutorías individuales o en pequeños grupos en el despacho del profesor, clase o laboratorio (TUT)
Estudio o preparación de pruebas [AUTÓNOMA]	Trabajo autónomo	CM05 CM7 INS01 SIS01 SIS03 SIS09	1.8	45	N	-	Estudio individual (EST)
Otra actividad no presencial [AUTÓNOMA]	Prácticas	CM05 CM7 INS01 INS04 PER02 PER04 PER05 SIS03 SIS09 UCLM03	0.9	22.5	N	-	Preparación de prácticas de laboratorio (PLAB)
Resolución de problemas o casos [PRESENCIAL]	Resolución de ejercicios y problemas	CM05 CM7 INS01 INS04 PER02 PER04 PER05 SIS01 SIS09	0.6	15	S	N	Resolución de ejercicios por parte del profesor y los estudiantes (PRO)
Elaboración de informes o trabajos [AUTÓNOMA]	Trabajo autónomo	CM05 CM7 INS01 INS04 INS05 PER02 PER04 PER05 SIS01 SIS03 SIS09 UCLM03	0.9	22.5	S	N	Realización de un informe sobre un tema propuesto por el profesor (RES)
Prácticas de laboratorio [PRESENCIAL]	Prácticas	CM05 CM7 INS04 PER02 PER04 PER05 SIS03 SIS09	0.72	18	S	S	Realización en el laboratorio de las prácticas programadas (LAB)
Pruebas de progreso [PRESENCIAL]	Pruebas de evaluación	CM05 CM7 INS01 INS04 INS05 PER02 SIS01 SIS09 UCLM03	0.1	2.5	S	N	Realización de la prueba de progreso 1 correspondiente al primer tercio del temario de la asignatura (EVA)
Pruebas de progreso [PRESENCIAL]	Pruebas de evaluación	CM05 CM7 INS01 INS04 INS05 PER02 SIS01 SIS09 UCLM03	0.1	2.5	S	N	Realización de la prueba de progreso 2 correspondiente a los dos primeros tercios del temario de la asignatura (EVA)
Pruebas de progreso [PRESENCIAL]	Pruebas de evaluación	CM05 CM7 INS01 INS04 INS05 PER02 SIS01 SIS09 UCLM03	0.1	2.5	S	N	Realización de la prueba de progreso 3 correspondiente al temario completo de la asignatura (EVA)
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
Pruebas de progreso	15.00%	0.00%	Prueba de progreso 2. Actividad no obligatoria y recuperable a realizar al final del segundo tercio del periodo docente
Pruebas de progreso	27.50%	0.00%	Prueba de progreso 3. Actividad no obligatoria y recuperable a realizar en el periodo sin docencia
Elaboración de trabajos teóricos	15.00%	15.00%	Actividad no obligatoria y recuperable a realizar antes del fin del periodo docente
Pruebas de progreso	7.50%	0.00%	Prueba de progreso 1. Actividad no obligatoria y recuperable a realizar al final del primer tercio del periodo docente
Realización de prácticas en laboratorio	25.00%	25.00%	Actividad obligatoria y recuperable a realizar en las sesiones de laboratorio
Presentación oral de temas	10.00%	10.00%	Actividad no obligatoria y recuperable. A realizar en las sesiones de teoría/laboratorio para los estudiantes de la modalidad continua. Los estudiantes de modalidad no continua serán evaluados de esta actividad a través de un sistema alternativo en la convocatoria ordinaria
Prueba final	0.00%	50.00%	Actividad obligatoria y recuperable a realizar en la fecha prevista para el examen final de la convocatoria ordinaria
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:

En las actividades obligatorias se debe obtener un mínimo de 4 sobre 10 para considerar la actividad superada y poder aprobar la asignatura. Una misma actividad evaluable obligatoria no podrá dividirse en partes eliminatorias ni establecerse notas mínimas en cada una de sus partes. En las actividades recuperables existe una prueba de evaluación alternativa en la convocatoria extraordinaria.

Las pruebas de progreso serán comunes para todos los grupos de teoría/laboratorio de la asignatura y serán calificadas por los profesores de la asignatura de forma horizontal, es decir, cada una de las partes de las pruebas de progreso serán evaluadas por el mismo profesor para todos los estudiantes.

El estudiante aprueba la asignatura si obtiene un mínimo de 50 puntos sobre 100 con las valoraciones de cada actividad de evaluación y supera todas las actividades obligatorias.

Para los estudiantes que no aprueben la asignatura en la convocatoria ordinaria, la calificación de las actividades superadas se conservará para la convocatoria extraordinaria. Si una actividad no es recuperable, su valoración se conservará para la convocatoria extraordinaria aunque no se haya superado. En el caso de actividades recuperables superadas, el estudiante podrá presentarse a la evaluación alternativa de esas actividades en la convocatoria extraordinaria y, en ese caso, la nota final de la actividad corresponderá a la última nota obtenida.

La calificación de las actividades superadas en cualquier convocatoria, exceptuando las pruebas de progreso, se conservará para el próximo curso académico a petición del estudiante siempre que ésta sea igual o superior a 5 y no se modifique las actividades formativas y los criterios de evaluación de la asignatura en el próximo curso académico.

La no comparecencia a la prueba de progreso 3 supondrá la calificación de "No presentado". Si el estudiante no ha superado alguna actividad de evaluación obligatoria, la nota final en la asignatura no puede superar el 4 sobre 10.

Evaluación no continua:

Los estudiantes pueden solicitar, al principio del cuatrimestre, acogerse a la modalidad de evaluación no continua. Del mismo modo, el estudiante podrá cambiarse a la modalidad de evaluación no continua siempre que no haya participado durante el periodo de impartición de clases en actividades evaluables que supongan en su conjunto al menos el 50% de la evaluación total de la asignatura. Si un estudiante ha alcanzado ese 50% de actividades evaluables o si, en cualquier caso, el periodo de clases hubiera finalizado, se considerará en evaluación continua sin posibilidad de cambiar de modalidad de evaluación.

Los estudiantes que se acogen a la modalidad de evaluación no continua serán calificados globalmente, en 2 convocatorias anuales, una ordinaria y otra extraordinaria, evaluándose el 100% de las competencias, a través de los sistemas de evaluación indicados en la columna "Evaluación no continua".

En la modalidad de evaluación "no continua" no existe la obligatoriedad de conservar la nota obtenida por el estudiante en las actividades o pruebas (de progreso o parciales) que haya realizado en modalidad de evaluación continua.

Particularidades de la convocatoria extraordinaria:

Se realizarán pruebas de evaluación para todas las actividades recuperables. Por la naturaleza de las pruebas de progreso en la convocatoria extraordinaria existirá una única prueba de progreso que engloba a las tres pruebas de progreso de la convocatoria ordinaria. La no comparecencia a esta prueba de progreso supondrá la calificación de "No presentado", salvo en el caso en que el estudiante conserve la nota de dicha prueba de la convocatoria ordinaria. En este último caso, la realización por parte del estudiante de cualquier otra actividad evaluable en la convocatoria extraordinaria supondrá la calificación numérica en el acta.

Particularidades de la convocatoria especial de finalización:

Mismas características que en la convocatoria extraordinaria.

9. SECUENCIA DE TRABAJO, CALENDARIO, HITOS IMPORTANTES E INVERSIÓN TEMPORAL	
No asignables a temas	
Horas	Suma horas
Comentarios generales sobre la planificación: La asignatura se imparte en tres sesiones semanales de 1,5 horas	

10. BIBLIOGRAFÍA, RECURSOS

Autor/es	Título/Enlace Web	Editorial	Población	ISBN	Año	Descripción
Jan Van der Plass	Python Data Science Handbook https://learning.oreilly.com/library/view/python-data-science/9781491912126/	O'Reilly		9781491912058	2016	
Igual, Laura, Seguí, Santi	Introduction to Data Science https://link.springer.com/book/10.1007%2F978-3-319-50017-1	Springer		9783319500171	2017	This accessible and classroom-tested textbook/reference presents an introduction to the fundamentals of the emerging and interdisciplinary field of data science. The coverage spans key concepts adopted from statistics and machine learning and the practical application of data science.
Adriaans, P. W.; Zantinge, D.	Data Mining.	Addison-Wesley			1996	
Berry, M. J. A.; Linoff, G.	Data Mining Techniques.	Wiley Computer Publishing.	New York		1996	
Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P.	The KDD Process for Extracting Useful Knowledge from Volumes of Data.				1996	
Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R. (Eds)	Advances in Knowledge Discovery and Data Mining.	AAAI/MIT Press	Cambridge	MA	1996	
Jefrey Leek	The Elements of Data Analytic Style http://worldpece.org/sites/default/files/datastyle.pdf	LeanPub			2014	