UNIVERSIDAD DE CASTILLA - LA MANCHA

GUÍA DOCENTE

1. DATOS GENERALES

Asignatura: MINERÍA DE DATOS
Código: 42348
Tipología: OPTATIVA
Créditos ECTS: 6
Grado: 407 - GRADO EN INGENIERÍA INFORMÁTICA (CR)_20
Curso académico: 2020-21
Centro: 108 - ESCUELA SUPERIOR DE INFORMATICA C. REAL
Grupo(s): 20
Curso: 4
Duración: Primer cuatrimestre
Lengua principal de impartición: Español
Segunda lengua:
Uso docente de otras lenguas:
English Friendly: S
Página web:
Bilingüe: N

Profesor: **JOSE ANGEL OLIVAS VARELA** - Grupo(s): **20**				
Edificio/Despacho	**Departamento**	**Teléfono**	**Correo electrónico**	**Horario de tutoría**
Fermín Caballero / 3.10	TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN	6476	joseangel.olivas@uclm.es	Disponible en https://esi.uclm.es/categories/profesorado-y-tutorias

Profesor: **FRANCISCO PASCUAL ROMERO CHICHARRO** - Grupo(s): **20**				
Edificio/Despacho	**Departamento**	**Teléfono**	**Correo electrónico**	**Horario de tutoría**
Fermin Caballero / 3.17	TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN	926051535	franciscop.romero@uclm.es	Disponible en https://esi.uclm.es/categories/profesorado-y-tutorias

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 lasasignaturas:

- 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 márketing, 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.
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

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

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.

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 CM07	0.6	15	N	-	Exposición del temario por parte del profesor (MAG)
Tutorías individuales [PRESENCIAL]		CM05 CM07 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 CM07 INS01 SIS01 SIS03 SIS09	1.8	45	N	-	Estudio individual (EST)
Otra actividad no presencial [AUTÓNOMA]	Prácticas	CM05 CM07 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 CM07 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 CM07 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 CM07 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 CM07 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 CM07 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 CM07 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	Evaluacion continua	Evaluación no continua*	Descripción
Prueba final	0.00%	50.00%	Actividad obligatoria y recuperable a realizar en la fecha prevista para el examen final de la convocatoria ordinaria
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
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
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.
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 valoración de las actividades será global y, por tanto, se debe expresar por medio de una única nota. Si la actividad consta de varios apartados podrá valorarse de forma individual informando por escrito durante el inicio del curso sobre los criterios de valoración de cada apartado.En las actividades obligatorias se debe obtener un mínimo de 4 sobre 10 para considerar la actividad superada y poder aprobar la asignatura. 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. 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 que no puedan asistir regularmente a las actividades formativas presenciales pueden solicitar, al principio del cuatrimestre, acogerse a la modalidad de evaluación no continua. Del mismo modo, si un estudiante que está realizando la modalidad de evaluación continua, incurre en alguna circunstancia que le impida asistir regularmente a las actividades formativas presenciales, puede acogerse a la modalidad de evaluación no continua. En este caso se debe notificar antes de la fecha prevista para las pruebas de la convocatoria ordinaria, de acuerdo con un plazo límite que se informará al inicio del semestre. 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 realizará 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.

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	O'Reilly		9781491912058	2016	
	https://learning.oreilly.com/library/view/python-data-science/9781491912126/					
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	LeanPub			2014	
	http://worldpece.org/sites/default/files/datastyle.pdf					
Igual, Laura, Seguí, Santi	Introduction to Data Science	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.
	https://link.springer.com/book/10.1007%2F978-3-319-50017-1					