

Escuela Superior de Ingeniería y Tecnología

Grado en Ingeniería Informática

GUÍA DOCENTE DE LA ASIGNATURA :

**Tratamiento Inteligente de Datos
(2022 - 2023)**

1. Datos descriptivos de la asignatura

Asignatura: Tratamiento Inteligente de Datos	Código: 139263125
- Centro: Escuela Superior de Ingeniería y Tecnología - Lugar de impartición: Escuela Superior de Ingeniería y Tecnología - Titulación: Grado en Ingeniería Informática - Plan de Estudios: 2010 (Publicado en 2011-03-21) - Rama de conocimiento: Ingeniería y Arquitectura - Itinerario / Intensificación: - Departamento/s: Ingeniería Informática y de Sistemas Ingeniería Industrial - Área/s de conocimiento: Ciencia de la Computación e Inteligencia Artificial Ingeniería Telemática Lenguajes y Sistemas Informáticos - Curso: 3 - Carácter: Obligatoria - Duración: Segundo cuatrimestre - Créditos ECTS: 6,0 - Modalidad de impartición: Presencial - Horario: Enlace al horario - Dirección web de la asignatura: http://www.campusvirtual.ull.es - Idioma: Español e Inglés	

2. Requisitos para cursar la asignatura

No existen requisitos para cursar la asignatura

3. Profesorado que imparte la asignatura

Profesor/a Coordinador/a: JOSE MARCOS MORENO VEGA
- Grupo: Grupos completo, PA y PE
General - Nombre: JOSE MARCOS - Apellido: MORENO VEGA - Departamento: Ingeniería Informática y de Sistemas - Área de conocimiento: Ciencia de la Computación e Inteligencia Artificial

Contacto

- Teléfono 1: **922318175**
- Teléfono 2:
- Correo electrónico: **jmmoreno@ull.es**
- Correo alternativo:
- Web: **<http://www.campusvirtual.ull.es>**

Tutorías primer cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Martes	08:00	10:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019
Todo el cuatrimestre		Viernes	08:00	12:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019

Observaciones:

Tutorías segundo cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Miércoles	08:00	10:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019
Todo el cuatrimestre		Viernes	08:00	12:00	Escuela Superior de Ingeniería y Tecnología - Módulo A - AN.4A ESIT	P3.019

Observaciones:

Profesor/a: PATRICIO GARCIA BAEZ

- Grupo: **Grupos completos, PA y PE**

General

- Nombre: **PATRICIO**
- Apellido: **GARCIA BAEZ**
- Departamento: **Ingeniería Informática y de Sistemas**
- Área de conocimiento: **Ciencia de la Computación e Inteligencia Artificial**

Contacto

- Teléfono 1: **922845038**
- Teléfono 2:
- Correo electrónico: **pgarcia@ull.es**
- Correo alternativo:
- Web: **<http://www.campusvirtual.ull.es>**

Tutorías primer cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Lunes	15:00	18:00	Escuela Superior de Ingeniería y Tecnología - Módulo C - AN.4A ESIT	P2.102
Todo el cuatrimestre		Miércoles	10:30	13:30	Escuela Superior de Ingeniería y Tecnología - Módulo C - AN.4A ESIT	P2.102

Observaciones: Consultar actualizaciones y modificaciones puntuales de este horario en perfil del profesor del Campus Virtual

Tutorías segundo cuatrimestre:

Desde	Hasta	Día	Hora inicial	Hora final	Localización	Despacho
Todo el cuatrimestre		Martes	15:00	18:00	Escuela Superior de Ingeniería y Tecnología - Módulo C - AN.4A ESIT	P2.102
Todo el cuatrimestre		Miércoles	10:00	13:00	Escuela Superior de Ingeniería y Tecnología - Módulo C - AN.4A ESIT	P2.102

Observaciones: Consultar actualizaciones y modificaciones puntuales de este horario en perfil del profesor del Campus Virtual

4. Contextualización de la asignatura en el plan de estudio

Bloque formativo al que pertenece la asignatura: **Itinerario 1: Computación**
Perfil profesional: **Ingeniero Técnico en Informática**

5. Competencias

Tecnología Específica / Itinerario: Computación

C45 - 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.

Competencias Generales

CG4 - Capacidad para definir, evaluar y seleccionar plataformas hardware y software para el desarrollo y la ejecución de sistemas, servicios y aplicaciones informáticas, de acuerdo con los conocimientos adquiridos según lo establecido en el apartado 5 de este anexo.

CG6 - Capacidad para concebir y desarrollar sistemas o arquitecturas informáticas centralizadas o distribuidas integrando hardware, software y redes de acuerdo con los conocimientos adquiridos según lo establecido en el apartado 5 de este anexo.

CG9 - Capacidad para resolver problemas con iniciativa, toma de decisiones, autonomía y creatividad. Capacidad para saber comunicar y transmitir los conocimientos, habilidades y destrezas de la profesión de Ingeniero Técnico en Informática.

Transversales

T7 - Capacidad de comunicación efectiva (en expresión y comprensión) oral y escrita, con especial énfasis en la redacción de documentación técnica.

T10 - Capacidad de integrarse rápidamente y trabajar eficientemente en equipos unidisciplinarios y de colaborar en un entorno multidisciplinar.

T21 - Capacidad para el razonamiento crítico, lógico y matemático.

T23 - Capacidad de abstracción: capacidad de crear y utilizar modelos que reflejen situaciones reales.

6. Contenidos de la asignatura

Contenidos teóricos y prácticos de la asignatura

Módulo: INTRODUCCIÓN

(Profesorado: J. Marcos Moreno Vega, Patricio García Báez (teoría y práctica))

- Tema 1. El proceso de extracción de conocimiento en grandes volúmenes de datos
- Tema 2. Recuperación de información en documentos electrónicos

Módulo: CLASIFICACIÓN, REGRESIÓN Y AGRUPAMIENTO

(Profesorado: J. Marcos Moreno Vega (teoría y práctica))

- Tema 3. Árboles de decisión y regresión
- Tema 4. Clasificadores bayesianos
- Tema 5. Agrupamiento basado en prototipos
- Tema 6. Agrupamiento basado en densidad
- Tema 7. Agrupamiento jerárquico

Módulo: PREPROCESADO DE DATOS

(Profesorado: J. Marcos Moreno Vega (teoría y práctica))

- Tema 8. Preparación de datos
- Tema 9. Reducción de la dimensionalidad

Módulo: CLASIFICACIÓN Y AGRUPAMIENTO MEDIANTE REDES NEURONALES

(Profesorado: Patricio García Báez (teoría y práctica))

- Tema 10. Clasificadores neuronales
- Tema 11. Agrupamiento mediante redes neuronales

Módulo: MINERÍA DE PATRONES DE ASOCIACIÓN

(Profesorado: J. Marcos Moreno Vega (teoría y práctica))

- Tema 12. Reglas de asociación

Actividades a desarrollar en otro idioma

El software utilizado en la prácticas de laboratorio está documentado en inglés. También lo están los manuales y tutoriales que usarán los alumnos para el desarrollo de las prácticas asociadas a esta asignatura. Además, gran parte del material bibliográfico y de trabajo de la asignatura está escrito en este idioma. Se refuerza con ello la comprensión lectora de este idioma por parte de los alumnos.

Por otro lado, la memoria del proyecto (que se describe en el apartado Metodología y volumen del trabajo del estudiante) incluirá un breve resumen en inglés y los alumnos expondrán oralmente las conclusiones de su trabajo en este idioma. La evaluación de estas actividades se incluyen en la evaluación del proyecto, según lo descrito en el apartado Sistema de evaluación y calificación.

7. Metodología y volumen de trabajo del estudiante

Descripción

Los contenidos de la asignatura serán explicados por los profesores en las horas semanales presenciales de teoría. Durante las clases prácticas, se explicará la herramienta de tratamiento de datos usada como apoyo a los contenidos teóricos. Esta herramienta será usada por los alumnos para realizar las actividades prácticas que así lo requieran.

Para fomentar la comprensión del proceso de extracción de conocimiento y de las técnicas asociadas, se usará la metodología de enseñanza aprendizaje basada en proyectos, junto con la realización de pruebas prácticas.

Proyecto.-

- Con la realización del proyecto se persigue favorecer el aprendizaje efectivo, potenciar el trabajo autónomo, reforzar la motivación e implicación y favorecer la actitud reflexiva y crítica de los alumnos.
- Se formarán grupos de alumnos que tendrán que recopilar, preparar y tratar datos con el propósito de extraer conocimiento útil de los mismos. Además, deberán interpretar los resultados obtenidos y proponer estrategias para difundir y usar el conocimiento extraído. Se procurará que los datos a analizar sean de interés para el alumnado. Así, se usarán datos sobre hábitos de compra, perfiles de usuarios de algún servicio o relativos a variables económicas o medioambientales.
- El proyecto se estructurará en dos subproyectos. En el primero se describirán y emplearán técnicas clásicas de análisis de datos, mientras que en el segundo se usarán técnicas basadas en redes neuronales.
- El trabajo realizado se recogerá en una memoria del proyecto) que deberá ser presentada oralmente al finalizar el cuatrimestre. El resumen del proyecto y las conclusiones derivadas del mismo se redactarán en inglés.
- La memoria escrita y la exposición se usarán como base para la evaluación continua.
- En dos semanas del cuatrimestre los alumnos expondrán a los profesores, en una tutoría en grupo, la marcha del proyecto. En estas tutorías los profesores evaluarán la marcha del proyecto y sugerirán, en su caso, acciones de mejora.

Pruebas prácticas.-

- Con estas actividades se pretende que el alumnado profundice en los fundamentos teóricos del análisis de datos y en su uso para afrontar supuestos prácticos reales.
- Se contempla la realización de cinco pruebas prácticas que tratarán sobre clasificación, agrupamiento y regresión con técnicas clásicas y mediante redes neuronales.

Otras actividades complementarias a desarrollar por el alumno son la participación en foros, la búsqueda de información reciente sobre el tratamiento de datos y sus implicaciones legales, éticas y morales y la realización de ejercicios o pruebas. Algunas de estas actividades se realizarán virtualmente usando para ello las herramientas (foros, cuestionarios ...) de la plataforma de apoyo a la docencia virtual de la ULL.

Actividades formativas en créditos ECTS, su metodología de enseñanza-aprendizaje y su relación con las competencias que debe adquirir el estudiante

Actividades formativas	Horas presenciales	Horas de trabajo autónomo	Total horas	Relación con competencias
Clases teóricas	16,00	0,00	16,0	[T23], [C45]

Clases prácticas (aula / sala de demostraciones / prácticas laboratorio)	14,00	0,00	14,0	[T23], [T21], [C45]
Realización de seminarios u otras actividades complementarias	8,00	10,00	18,0	[T21], [T10], [T7]
Realización de trabajos (individual/grupal)	7,00	40,00	47,0	[T23], [T21], [T10], [T7], [CG9], [CG6], [CG4], [C45]
Estudio/preparación de clases teóricas	0,00	10,00	10,0	[T23], [T21], [C45]
Realización de exámenes	3,00	0,00	3,0	[T7], [CG9], [CG6], [CG4]
Asistencia a tutorías	7,00	0,00	7,0	[C45]
Estudio autónomo individual o en grupo	0,00	30,00	30,0	[T23], [T21], [T10], [CG9], [CG6], [CG4]
Exposición oral por parte del alumno	5,00	0,00	5,0	[T10], [T7]
Total horas	60,00	90,00	150,00	
Total ECTS			6,00	

8. Bibliografía / Recursos

Bibliografía Básica

José Hernández Orallo, M.José Ramírez Quintana, Cesar Ferri Ramírez. Introducción a la Minería de Datos Editorial Pearson, 2004.
 Ethem Alpaydin, Introduction to Machine Learning, The MIT Press, 2004
 Ian H. Witten, Eibe Frank, Mark A. Hall, Data Mining, Morgan Kaufmann, 2011
 P. N. Tan, M. Steinbach, V. Kumar. Introduction to Data Mining, Addison-Wesley, 2006
 C. C. Data Mining. The Textbook. Springer, 2015
 S. García, J. Luengo, F. Herrera. Data preprocessing in Data Mining. Springer. 2015

Bibliografía Complementaria

Tom Mitchell, Machine Learning, McGraw Hill, 1997
 Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009

Otros Recursos

9. Sistema de evaluación y calificación

Descripción

La evaluación de la asignatura se rige por el Reglamento de Evaluación y Calificación de la Universidad de La Laguna, aprobado el 21 de junio de 2022, o el que la Universidad tenga vigente, además de por lo establecido en la Memoria de Verificación inicial o posteriores modificaciones.

Se emplea un sistema de evaluación continua que tiene en cuenta los siguientes actividades relacionadas con las pruebas descritas en el apartado "Metodología y volumen de trabajo":

- Trabajo y defensa del subproyecto 1: técnicas clásicas de análisis de datos.
- Trabajo y defensa del subproyecto 2: técnicas basadas en redes neuronales.
- Informe del subproyecto 1: técnicas clásicas de análisis de datos.
- Informe del subproyecto 2: técnicas basadas en redes neuronales.
- Pruebas prácticas.

Cada subproyecto tiene un peso del 25% en la nota final (10% corresponde al informe y 15% al trabajo y su defensa), correspondiendo el 50% restante a cinco pruebas prácticas. Cada una de estas pruebas tendrá un peso del 10% en la nota final.

Todas las actividades de la evaluación continua tienen carácter obligatorio.

En el apartado Cronograma/calendario de la asignatura se recogen las fechas estimadas de presentación de las diferentes actividades. Los resultados de las mismas serán comunicados a los alumnos aproximadamente 15 días después de su presentación.

La evaluación continua, obligatoria en la primera convocatoria de la asignatura para todo el alumnado (excepto para quienes se acojan a la evaluación única según se dispone en el artículo 5.4 del Reglamento de Evaluación y Calificación de la Universidad de La Laguna) se mantendrá en la segunda convocatoria. El alumnado que no haya superado la asignatura en la primera convocatoria deberá examinarse de todas las actividades obligatorias que no haya superado en dicha convocatoria. En su caso, la entrega y defensa oral del proyecto y la realización de las pruebas prácticas se realizarán en las fechas establecidas para ello por la Universidad de La Laguna y la Escuela Superior de Ingeniería y Tecnología.

La evaluación única, recogida en el Reglamento de Evaluación y Calificación de la Universidad de La Laguna, constará de ejercicios teóricos y prácticos. La prueba se celebrará en las fechas establecidas para ello por la Universidad de La Laguna y la Escuela Superior de Ingeniería y Tecnología.

Tal como se ha indicado anteriormente, el software utilizado en la prácticas de laboratorio, así como los manuales y tutoriales que usarán los alumnos para el desarrollo de las mismas, están documentados en inglés. Se evaluará la comprensión de los mismos en la calificación de prácticas.

En la tabla siguiente se muestra la ponderación de las diferentes actividades de la evaluación continua.

Estrategia Evaluativa

Tipo de prueba	Competencias	Criterios	Ponderación
Valoración de las actividades prácticas en el laboratorio	[T23], [T21], [T10], [T7], [CG9], [CG6], [CG4], [C45]	- Grado de conocimiento adquirido en el manejo de las herramientas para el tratamiento inteligente de datos. - Rigor en el análisis de los datos y coherencia de las conclusiones.	50,00 %
Elaboración de informes	[CG4], [CG6], [T7], [T10], [T23], [T21], [CG9], [C45]	- Memoria/s: estructura calidad y claridad de la redacción, fuentes consultadas, rigor en el análisis de los datos y coherencia de las conclusiones. - Defensa oral: estructura de la exposición, lenguaje empleado y respuesta a las preguntas.	20,00 %
Realización de trabajos y su defensa y/o exposición	[CG4], [CG6], [T7], [T23], [T21], [CG9], [C45]	Adecuación de las respuestas dadas a los ejercicios y claridad de redacción de las mismas.	30,00 %

10. Resultados de Aprendizaje

Diseñar, implementar y evaluar técnicas de aprendizaje computacional y de extracción automática de conocimiento a partir de grandes volúmenes de datos.

En el contexto de un grupo de trabajo, recopilar, analizar y estructurar documentación técnica sobre los aspectos claves de las técnicas de aprendizaje computacional y extracción automática de información y conocimiento a partir de grandes volúmenes de datos para realizar una síntesis por escrito y una presentación oral

11. Cronograma / calendario de la asignatura

Descripción

Los contenidos de la asignatura serán explicados por el profesorado en las horas semanales presenciales de teoría. Para afianzar estos contenidos, se han programado diferentes tareas prácticas que los alumnos deberán hacer individualmente y de forma autónoma. En el enunciado de estas tareas se especificarán los criterios de evaluación y, en su caso, la fecha límite de entrega.

Se pondrá a disposición de los alumnos las transparencias de clase, ejemplos y enlaces a material complementario.

Además, los alumnos deberán participar en dos foros online dedicados, respectivamente, a los "Límites del tratamiento inteligente de datos" e "Importancia, alcance, retos y oportunidades del tratamiento de datos". El profesorado enlazará en el aula virtual lecturas y material recomendados para que los alumnos puedan preparar adecuadamente sus intervenciones en el foro.

La asignatura sigue una metodología de enseñanza aprendizaje basada en proyectos por lo que los alumnos deberán realizar, en grupo, un proyecto en el que tendrán que recopilar, preparar y tratar datos con el propósito de extraer conocimiento útil de los mismos. Además, deberán interpretar los resultados obtenidos y proponer estrategias para difundir y usar el conocimiento extraído. Semanalmente se irá trabajando en este proyecto que deberá ser expuesto oralmente al finalizar el cuatrimestre. Las semanas 6 y 11 los alumnos presentarán al profesorado, en un seminario en grupo reducido, la marcha del proyecto. En estos seminarios el profesorado evaluará la marcha del proyecto y sugerirá, en su caso, acciones de mejora.

El cronograma que se muestra tiene carácter orientativo, de modo que el profesorado podrá modificar la planificación propuesta si así fuese necesario para una correcta marcha del proceso de enseñanza-aprendizaje.

Segundo cuatrimestre					
Semana	Temas	Actividades de enseñanza aprendizaje	Horas de trabajo presencial	Horas de trabajo autónomo	Total
Semana 1:	Tema 1	Explicar el tema 1 (El proceso de extracción de conocimiento en bases de datos) Explicar las principales funcionalidades y entornos del software de análisis de datos Participar en el foro Límites del tratamiento de datos (Actividad online)	4.00	3.00	7.00
Semana 2:	Tema 2	Explicar el tema 2 (Recuperación de información en documentos electrónicos). Entrega o realización de la prueba práctica: Recuperación de información en la Web Describir el proyecto que los alumnos deben realizar en grupo	4.00	0.00	4.00
Semana 3:	Tema 3	Explicar el tema 3 (Árboles de decisión y regresión) Ejercicios y práctica: Árboles de decisión y regresión Desarrollo del proyecto	4.00	6.00	10.00
Semana 4:	Tema 4	Explicar el tema 4 (Clasificadores bayesianos) Ejercicios y práctica: Clasificadores bayesianos Desarrollo del proyecto	4.00	5.00	9.00

Semana 5:	Tema 5	Explicar el tema 5 (Agrupamiento basado en prototipos) Ejercicios y práctica: Agrupamiento basado en prototipos Entrega o realización de la prueba práctica: Clasificación con técnicas clásicas Desarrollo del proyecto Participar en la tutoría en grupo reducido "Avances del proyecto"	4.00	4.00	8.00
Semana 6:	Tema 6	Explicar el tema 6 (Agrupamiento basado en densidad) Ejercicios y práctica: Agrupamiento basado en densidad Desarrollo del proyecto	4.00	6.00	10.00
Semana 7:	Tema 7	Explicar el tema 7 (Agrupamiento jerárquico) Ejercicios y práctica: Agrupamiento jerárquico Desarrollo del proyecto	4.00	6.00	10.00
Semana 8:	Tema 8	Explicar el tema 8 (Preparación de datos) Entrega o realización de la prueba práctica: Agrupamiento con técnicas clásicas Desarrollo del proyecto	4.00	5.00	9.00
Semana 9:	Tema 9	Explicar el tema 9 (Reducción de la dimensionalidad) Participar en la tutoría en grupo reducido "Avances del proyecto" Desarrollo del proyecto	4.00	6.00	10.00

Semana 10:	Tema 10	Explicar el tema 10 (Clasificadores neuronales) Entrega o realización de la prueba práctica: Clasificadores neuronales Desarrollo del proyecto	4.00	5.00	9.00
Semana 11:	Tema 10	Explicar el tema 10 (Clasificadores neuronales) Desarrollo del proyecto Participar en la tutoría en grupo reducido "Avances del proyecto"	4.00	6.00	10.00
Semana 12:	Tema 11	Explicar el tema 11 (Agrupamiento mediante redes neuronales) Entrega o realización de la prueba práctica: Agrupamiento mediante redes neuronales Desarrollo del proyecto	4.00	6.00	10.00
Semana 13:	Tema 12	Explicar el tema 12 (Reglas de asociación) Ejercicios y práctica: Reglas de asociación Desarrollo del proyecto	4.00	5.00	9.00
Semana 14:	Presentación oral del proyecto		4.00	6.00	10.00
Semana 15:	Evaluación única y de recuperación de las actividades obligatorias necesarias para superar la asignatura (semanas 15 y 16)	Preparación y realización de las pruebas evaluativas (redacción de la memoria, presentación del proyecto, prueba única (en su caso))	4.00	21.00	25.00
Semana 16 a 18:			0.00	0.00	0.00
Total			60.00	90.00	150.00