



FICHA TÉCNICA DE LA ASIGNATURA

| Datos de la asignatura | |
|------------------------|--------------------------------------------------------------------------|
| Nombre completo | Adquisición de datos |
| Código | DTC-IMAT-211 |
| Título | Grado en Ingeniería Matemática e Inteligencia Artificial |
| Impartido en | Grado en Ingeniería Matemática e Inteligencia Artificial [Segundo Curso] |
| Nivel | Reglada Grado Europeo |
| Cuatrimestre | Semestral |
| Créditos | 4,5 ECTS |
| Carácter | Obligatoria (Grado) |
| Departamento / Área | Departamento de Telemática y Computación |

| Datos del profesorado | |
|-----------------------|------------------------------------------|
| Profesor | |
| Nombre | Francisco Gómez Martín |
| Correo electrónico | fgmartin@icai.comillas.edu |
| Profesor | |
| Nombre | Ignacio Villanueva Romero |
| Departamento / Área | Departamento de Telemática y Computación |
| Correo electrónico | ivillanueva@icai.comillas.edu |

DATOS ESPECÍFICOS DE LA ASIGNATURA

| Contextualización de la asignatura |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Aportación al perfil profesional de la titulación |
| <p>Esta asignatura forma parte del bloque de asignaturas metodologías dentro del área de "Computación" junto con Visualización y Bases de datos. En este curso, conocerás los fundamentos del trabajo con datos, entendiendo la obtención de las diferentes fuentes, la limpieza y organización de los datos y su posterior uso, todo ello a través de bloques prácticos con diversos proyectos orientados a afianzar los conocimientos y poder extrapolarlos al día a día de la empresa.</p> |
| Prerrequisitos |
| Programación en python. |

| Competencias - Objetivos |
|--------------------------|
| Competencias |



GENERALES

| | |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CG04 | Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería. |
| CG05 | Conocimiento de la estructura, organización, funcionamiento e interconexión de los sistemas informáticos, los fundamentos de su programación, y su aplicación para la resolución de problemas propios de la ingeniería |
| CG07 | Capacidad para integrarse en equipos de trabajo y colaborar de forma activa con otras personas, áreas y organizaciones en la consecución de los objetivos ligados a las actividades de extracción de valor de los datos e inteligencia artificial. |

ESPECÍFICAS

| | |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CE14 | Dominio de los conceptos y técnicas más utilizadas de adquisición y transformación de la información localizada en local o en remoto en el ámbito del análisis de datos y la inteligencia artificial |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Resultados de Aprendizaje

| | |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| RA1 | Conocer los procedimientos y fases para implementar correctamente un sistema de adquisición de datos |
| RA2 | Implementar soluciones de ingesta de datos de cualquier fuente y naturaleza para su análisis posterior |
| RA3 | Extraer, procesar y transformar los datos con el objetivo de poseer un dato de calidad: limpio, homogéneo y estandarizado |
| RA4 | Extraer de forma programática la información deseada de los formatos de ficheros de datos más utilizados en la actualidad (TXT, CSV, XML, JSON, ..) |
| RA5 | Conocer el sistema de etiquetado del lenguaje HTML con el objetivo de identificar la información relevante dentro de las páginas web |
| RA6 | Dominar el proceso de automatización y extracción de información de páginas web |

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Bloque 1 Extracción y transformación de datos

- Teoría de autómatas
- Expresiones regulares

Bloque 2. Limpieza y calidad de los datos

- Integridad y calidad de los datos
- Limpieza y normalización
 - Gestión de codificaciones, manipulación de fechas, ...
- Imputación de datos

Bloque 3. Organización de los datos



- Descripción de los principales tipos de archivo (txt, JSON, XML, CSV)
- Procesos de conversión

Bloque 4. Ficheros de almacenamiento de información

- Análisis de distintos tipos de fichero (PDF, XLS, DOC)
- Procesos de importación/exportación

Bloque 5. La web como fuente de información

- Lenguaje HTML: etiquetas y estructura DOM
- Herramientas y librerías de webscraping
- Introducción a HTTP
- Introducción a APIs
- Automatización de procesos de obtención de datos desde la web

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

Las actividades formativas serán:

- **Clases magistrales expositivas y participativas:**
 - El profesor combinará exposición de los contenidos teóricos y con ejemplos prácticos, tanto matemáticos como de programación.
 - El alumno dispondrá de algunos ejemplos prácticos de código, generado dentro y fuera del aula por el profesor.
- **Ejercicios prácticos y resolución de problemas:**
 - El alumno resolverá problemas planteados por el profesor de forma presencial durante la clase, fomentando las dinámicas de trabajo cooperativo.
- **Sesiones prácticas con uso de software:**
 - Se dedicarán las sesiones prácticas a resolver dudas de las practicas relativas a cada bloque.
 - Las prácticas se graduarán por dificultad que los alumnos irán implementando conforme vayan finalizando cada hito.
- **Actividades de evaluación continua del rendimiento:** se realizarán pruebas relacionadas con las practicas semanales, como por ejemplo la exposición de resultados y la defensa de las soluciones propuestas frente a otras alternativas.

CG04, CG05, CG07, CE14

Metodología No presencial: Actividades

Las actividades formativas serán:

- **Ejercicios prácticos y resolución de problemas:**
 - El alumno dispondrá de problemas/ejemplos concretos enfocados a asimilar los conceptos explicados teóricos en las sesiones de teoría para desarrollar de forma no presencial.
- **Sesiones prácticas con uso de software:**



| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| <ul style="list-style-type: none"> ◦ Una vez liberada la práctica semanal después de la sesión de teoría correspondiente, el alumno trabajará sobre ella de forma no presencial. • Estudio personal: el objetivo principal del trabajo no presencial es llegar a entender y comprender los conceptos teóricos de la asignatura, así como ser capaz de poner en práctica estos conocimientos para resolver los diferentes tipos de problemas. Después de cada explicación teórica el profesor subirá a la web todos los códigos desarrollados y el alumno deberá revisarlos y plantearse cuestiones "Whatif" para asimilar mejor los conceptos teóricos. | CG04, CG05, CG07, CE14 |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|

RESUMEN HORAS DE TRABAJO DEL ALUMNO

| HORAS PRESENCIALES | | | | |
|-------------------------------------------------|----------------------------------------|------------------------------------------------|------------------------------------------------|----------------------------------------------------|
| Clases magistrales expositivas y participativas | Sesiones prácticas con uso de software | Ejercicios prácticos y resolución de problemas | Tutorías para resolución de dudas | Actividades de evaluación continua del rendimiento |
| 24.00 | 10.00 | 4.00 | 5.00 | 2.00 |
| HORAS NO PRESENCIALES | | | | |
| Sesiones prácticas con uso de software | Estudio personal | Proyectos | Ejercicios prácticos y resolución de problemas | |
| 40.00 | 15.00 | 30.00 | 4.00 | |
| CRÉDITOS ECTS: 4,5 (134,00 horas) | | | | |

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

| Actividades de evaluación | Criterios de evaluación | Peso |
|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Exámenes: <ul style="list-style-type: none"> • Prueba Intersemestral (15%) • Examen Final (50%) | <ul style="list-style-type: none"> • Prueba Intersemestral (15%): evaluación de los conocimientos adquiridos en teoría de autómatas, expresiones regulares y limpieza de datos mediante resolución de problemas gracias a la programación. • Examen Final (50%): se evaluará el pensamiento computacional y abstracto para la resolución de problemas de extracción, limpieza, organización y almacenamiento de datos mediante la programación. | 65 % |
| Practicas relacionadas con cada bloque (15%) | Se evaluarán regularmente los conocimientos adquiridos mediante un caso práctico individual que se deberá resolver mediante programación. | 15 % |
| Trabajo Final de la Asignatura (20%) | Se realizará un trabajo final que podrá ser tanto individual como en grupo y englobara una gran parte de los conocimientos adquiridos. | 20 % |



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

GUÍA DOCENTE

2024 - 2025

La calificación final en **convocatoria ordinaria** y **extraordinaria** de la asignatura dependerá de la evaluación de las siguientes actividades:

Nota Final = 15% Prueba_Intersemestral + 50% Examen_Final + 20% Proyecto final + 15% Casos prácticos

Sólo se aplicará esta nota final si se obtiene una **calificación mínima** de 5.0 en el Examen_Final.

La inasistencia al 15% o más de las horas presenciales de esta asignatura puede tener como consecuencia la imposibilidad de presentarse a las convocatorias ordinaria y extraordinaria.

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Presentaciones y códigos proporcionados por los profesores de la asignatura.

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

[https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792](https://servicios.upcomillas.es/sedeelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792)