

1. DATOS BÁSICOS

Asignatura	Procesamiento de datos
Titulación	Máster Universitario en Análisis de Datos Masivos (Big Data)
Escuela/ Facultad	Escuela de Arquitectura, Ingeniería y Diseño
Curso	1
ECTS	6
Carácter	Obligatoria
Idioma/s	Español
Modalidad	Presencial / A distancia
Semestre	1, 2
Curso académico	2023-2024
Docente coordinador	Nicolas Perez Coca

2. PRESENTACIÓN

Esta asignatura obligatoria de 6 ECTS prepara al estudiante para trabajar en proyectos de análisis de datos utilizando grandes volúmenes de datos. El estudiante utilizará los fundamentos de programación en Python necesarios para profundizar en el uso de este lenguaje en distintas tareas de análisis y procesamiento, así como el uso de herramientas de referencia usadas en la industria. El alumno se formará en la ingesta y el tratamiento de datos, y adquirirá conocimientos de todo el ciclo de vida del preprocesado de datos para el aprendizaje automático, desde la ingesta a el procesado y preparación de los mismos hasta algunos ejemplos de su uso con modelos de aprendizaje automático.

3. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

Las competencias que se desarrollan en esta asignatura son las siguientes:

- **Competencias Básicas:**
 - CB1. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
 - CB2. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
- **Competencias Transversales:**

CT1. Aprendizaje Autónomo: Habilidad para elegir las estrategias, las herramientas y los momentos que considere más efectivos para aprender y poner en práctica de manera independiente lo que ha aprendido.

CT5. Capacidad para aplicar los conocimientos a la práctica, para utilizar los conocimientos adquiridos en el ámbito académico en situaciones lo más parecidas posibles a la realidad de la profesión para la cual se están formando

CT10. Razonamiento crítico: Capacidad para analizar una idea, fenómeno o situación desde diferentes perspectivas y asumir ante él/ella un enfoque propio y personal, construido desde el rigor y la objetividad argumentada, y no desde la intuición.

- **Competencias Específicas:**

CE4. Diseñar y ejecutar un proceso completo de descubrimiento de conocimiento, incluyendo las fases de almacenamiento, procesamiento y visualización de los datos.

CE5. Diseñar y aplicar algoritmos de análisis basados en sistemas e infraestructuras de almacenamiento y acceso a grandes volúmenes de datos.

CE8. Aplicar las bases técnicas del funcionamiento de plataformas cloud computing y virtualizadas.

Los resultados de aprendizaje que se alcanzan en esta asignatura son los siguientes:

RA1: Diseñar, y aplicar algoritmos de minería de datos y aprendizaje automático para realizar un conjunto de test/pruebas y validar modelos

RA2: Utilizar modelos de extracción de patrones en grandes volúmenes de datos

En la tabla inferior se muestra la relación entre las competencias que se desarrollan en la asignatura y los resultados de aprendizaje que se persiguen:

Competencias	Resultados de aprendizaje
CB1, CB2, CT1, CT5, CT10, CE4, CE5	RA1
CB2, CB3, CT1, CT5, CT10, CE8	RA2

4. CONTENIDOS

Unidad 1. Introducción al análisis de datos

- Introducción al análisis de datos
- Los datos en el mundo real
- El formato de los datos
- Preparación de datos como proceso
- Etapas en la preparación del dato

Unidad 2. Ingesta de datos

- Descubrimiento y caracterización

- Muestreo y variabilidad
- Confianza
- Ingesta de datos en Pandas

Unidad 3. Preparación de los datos

- Valores faltantes
- Anomalías y Ruido
- Normalización y transformación de los datos
- Preparación de Series

Unidad 4. Análisis exploratorio de los datos

- Visualización analítica
- Dimensionalidad y muestreo del dataset
- Anonimización de datos

Unidad 5. Adecuación de los atributos

- Ingeniería de atributos (I)
- Ingeniería de atributos (II)
- Importancia de atributos
- Selección de atributos

Unidad 6. Análisis de documentos y textos

- Introducción al análisis de texto
- Representación y acceso al texto
- Preprocesamiento de documentos

5. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE

A continuación, se indican los tipos de metodologías de enseñanza-aprendizaje que se aplicarán:

- MD1: Clase magistral presencial o en su caso a través de campus virtual
- MD2: Aprendizaje Cooperativo
- MD3: Aprendizaje Basado en Problemas
- MD4: Investigación por grupos
- MD5: Aprendizaje basado en proyectos
- MD6: Entornos de simulación y experiencias de campo

6. ACTIVIDADES FORMATIVAS

A continuación, se identifican los tipos de actividades formativas que se realizarán y la dedicación en horas del estudiante a cada una de ellas:

Modalidad presencial:

Actividad formativa	Número de horas
Clases magistrales	30
Actividades individuales	35
Actividades aplicativas de carácter integrador	15
Trabajo autónomo	50
Tutorías, seguimiento	18
Pruebas de conocimiento	2

Modalidad online:

Actividad formativa	Número de horas
Clases magistrales	30
Actividades individuales	35
Actividades aplicativas de carácter integrador	15
Trabajo autónomo	50
Tutorías, seguimiento	18
Pruebas de conocimiento	2

7. EVALUACIÓN

A continuación, se relacionan los sistemas de evaluación, así como su peso sobre la calificación total de la asignatura:

Modalidad presencial:

Sistema de evaluación	Peso
SE1: Exámenes y tests.	45%
SE2: Elaboración de artículos, informes o memorias de diseños	35%
SE3: Evaluación entre compañeros de las simulaciones y de los casos prácticos.	20%

Modalidad online:

Sistema de evaluación	Peso
-----------------------	------

SE1: Exámenes y tests.	60%
SE2: Elaboración de artículos, informes o memorias de diseños	25%
SE3: Evaluación entre compañeros de las simulaciones y de los casos prácticos.	15%

En el Campus Virtual, cuando accedas a la asignatura, podrás consultar en detalle las actividades de evaluación que debes realizar, así como las fechas de entrega y los procedimientos de evaluación de cada una de ellas.

7.1. Convocatoria ordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de actividades.

7.2. Convocatoria extraordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de actividades.

Se deben entregar las actividades no superadas en convocatoria ordinaria, tras haber recibido las correcciones correspondientes a las mismas por parte del docente, o bien aquellas que no fueron entregadas.

8. CRONOGRAMA

En este apartado se indica el cronograma con fechas de entrega de actividades evaluables de la asignatura:

Actividades evaluables	Fecha (a partir de comienzo de asignatura)
Actividad 1	Semana 4
Actividad 2	Semana 6
Actividad 3	Semana 8
Actividad 4	Semana 10
Actividad 5	Semana 12
Examen final ordinaria	Semana 12

Este cronograma podrá sufrir modificaciones por razones logísticas de las actividades. Cualquier modificación será notificada al estudiante en tiempo y forma.

9. BIBLIOGRAFÍA

- García, S. Herrera, F. Luengo, J. (2015). Data Preprocessing in Data Mining. Springer.
- McKinney, W. (2017). Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython. USA. O'Reilly.
- Kuhn, M. Johnson, K. (2019). Feature Engineering and Selection: A Practical Approach for Predictive Models. Chapman & Hall.
- Kuhn, M. Johnson, K. (2013). Applied Predictive Modeling. Springer.
- James, G. Witten D. Hastie T. y Tibshirani, R. (2017) Introduction to Statistical Learning: with applications in R. Springer.
- Towards data science. <https://towardsdatascience.com/>
- Saul Dobilas. <https://solclover.com/>
- Python.org. <https://docs.python.org/es/3.8/tutorial/index.html>

10. UNIDAD DE ORIENTACIÓN EDUCATIVA Y DIVERSIDAD

Desde la Unidad de Orientación Educativa y Diversidad (ODI) ofrecemos acompañamiento a nuestros estudiantes a lo largo de su vida universitaria para ayudarles a alcanzar sus logros académicos. Otros de los pilares de nuestra actuación son la inclusión del estudiante con necesidades específicas de apoyo educativo, la accesibilidad universal en los distintos campus de la universidad y la equiparación de oportunidades.

Desde esta Unidad se ofrece a los estudiantes:

1. Acompañamiento y seguimiento mediante la realización de asesorías y planes personalizados a estudiantes que necesitan mejorar su rendimiento académico.
2. En materia de atención a la diversidad, se realizan ajustes curriculares no significativos, es decir, a nivel de metodología y evaluación, en aquellos alumnos con necesidades específicas de apoyo educativo persiguiendo con ello una equidad de oportunidades para todos los estudiantes.
3. Ofrecemos a los estudiantes diferentes recursos formativos extracurriculares para desarrollar diversas competencias que les enriquecerán en su desarrollo personal y profesional.
4. Orientación vocacional mediante la dotación de herramientas y asesorías a estudiantes con dudas vocacionales o que creen que se han equivocado en la elección de la titulación.

Los estudiantes que necesiten apoyo educativo pueden escribirnos a:
orientacioneducativa@universidadeuropea.es

11. ENCUESTAS DE SATISFACCIÓN

¡Tú opinión importa!

La Universidad Europea te anima a participar en las encuestas de satisfacción para detectar puntos fuertes y áreas de mejora sobre el profesorado, la titulación y el proceso de enseñanza-aprendizaje.

Las encuestas estarán disponibles en el espacio de encuestas de tu campus virtual o a través de tu correo electrónico.

Tu valoración es necesaria para mejorar la calidad de la titulación.

Muchas gracias por tu participación.