

## 1. DATOS BÁSICOS

<b>Asignatura</b>	MÓDULO 4 – ANÁLISIS Y APRENDIZAJE
<b>Titulación</b>	MÁSTER UNIVERSITARIO EN ANÁLISIS DE GRANDES CANTIDADES DE DATOS-MBI / BIG DATA ANALYTICS-MBI
<b>Escuela/ Facultad</b>	ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO
<b>Curso</b>	1
<b>ECTS</b>	12
<b>Carácter</b>	OBLIGATORIO
<b>Idioma/s</b>	ESPAÑOL
<b>Modalidad</b>	ONLINE / PRESENCIAL
<b>Semestre</b>	1
<b>Curso académico</b>	20/21
<b>Docente coordinador</b>	Enrique Puertas (Presencial) / José Javier Ruiz (Online)

## 2. PRESENTACIÓN

Módulo de 12 créditos ECTS (300 horas) que prepara al estudiante para trabajar en proyectos de análisis de datos, y enseña a construir modelos basados en Aprendizaje Automático utilizando grandes volúmenes de datos. El módulo se compone de dos partes complementarias: La asignatura de Análisis y Tecnologías, que enseña al estudiante los fundamentos de programación en Python necesarios para luego 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, y la asignatura de Aprendizaje Automático en la que el estudiante aprenderá a construir modelos de predicción, clasificación o agrupación (entre otros) a partir de grandes cantidades de datos

## 3. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

Los objetivos de aprendizaje son:

- RA1: Capacidad para diseñar, y aplicar algoritmos de minería de datos y aprendizaje automático para realizar un conjunto de test/pruebas y validar modelos
- RA2: Utilización de modelos de extracción de patrones en grandes volúmenes de datos
- RA3: Utilización de modelos de clasificación y predicción en grandes volúmenes de datos
- RA4: Utilización de modelos de aprendizaje en grandes volúmenes de datos

Las competencias básicas que se desarrollarán en la materia son:

- 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
- CB5: Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Las competencias transversales que se desarrollarán en la materia son:

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.

CT4: Capacidad de análisis y síntesis: ser capaz de descomponer situaciones complejas en sus partes constituyentes; también evaluar otras alternativas y perspectivas para encontrar soluciones óptimas. La síntesis busca reducir la complejidad con el fin de entenderla mejor y/o resolver problemas.

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.

CT11: Resolución de problemas: Capacidad de encontrar solución a una cuestión confusa o a una situación complicada sin solución predefinida, que dificulte la consecución de un fin

CT14: Trabajo en equipo: Capacidad para integrarse y colaborar de forma activa con otras personas, áreas y/u organizaciones para la consecución de objetivos comunes

Las competencias específicas que se desarrollarán en la materia son:

CE2. Conocimiento de las bases teórico-prácticas necesarias sobre Tecnologías de la Información y Comunicaciones de interés para el desarrollo e implantación de servicios de análisis y extracción de modelos a partir de los datos en infraestructuras de altas prestaciones

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

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

CE7. Conocimiento de los aspectos relativos a la integración, implantación y explotación de aplicaciones de análisis de datos en plataformas de altas prestaciones incluyendo la privacidad y protección de los datos.

CE13. Capacidad para identificar futuras tendencias técnicas en tecnologías y procesos de descubrimiento de información y generación de conocimiento a partir de los 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
CB2, CB4, CB5, CT1, CT4, CT5, CT10, CT11, CT14, CE2, CE4, CE5, CE7, CE13	RA1
CB1, CB2, CB4, CB5, CT5, CT11, CT14, CE2, CE7	RA2, RA3, RA4

## 4. CONTENIDOS

Este módulo se compone de dos materias principales.

- Procesamiento de datos:
  - o Minería de datos
  - o Transformación de los datos
    - ▣ Extracción y selección de características
    - ▣ Discretización, muestreo y normalización de los datos
    - ▣ Reducción de dimensionalidad
  - o Análisis de contenidos multimedia
  - o Modelos de predicción y probabilísticos
- Aprendizaje automático
  - o Aprendizaje supervisado y no supervisado
  - o Técnicas de aprendizaje y programación en paralelo
  - o Análisis de datos en tiempo real
- Tecnologías para el procesamiento de grandes cantidades de datos
  - o Open source (p.e. Apache Mahout)
  - o Propietarias (p.e. Rapidminer)

## 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 Seminario Virtual. Se utilizará para impartir los diferentes contenidos de cada materia y la realización de coloquios y debates.
- MD2-Aprendizaje cooperativo. Se utilizará para el análisis de casos promocionando la participación y el intercambio de conocimiento para la consecución de resultados.
- MD3-Aprendizaje basado en problemas (ABP). Se utilizará para la resolución de problemas los cuales permitirán bajo el paradigma de aprender haciendo aplicar los conocimientos adquiridos en las clases magistrales a casos reales o simulados
- MD5-Aprendizaje basado en proyectos. Se utilizarán proyectos definidos y reales/simulados para desarrollar habilidades de diseño, estratégicas y procedimentales que engloben más de una única materia perteneciente al máster.

- MD6-Entornos de simulación y experiencias de campo. Se utilizarán casos reales y en su caso dominios conocidos para el desarrollo del conocimiento condicionado al dominio experto y de trabajo donde se vayan a aplicar los conocimientos adquiridos.

## 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
Lecciones magistrales	50
Actividades individuales	50
Actividades aplicativas de carácter integrador	50
Trabajo autónomo	100
Tutorías, seguimiento y evaluación	50

### Modalidad online:

Actividad formativa	Número de horas
Seminario virtual	10
Actividades individuales	70
Actividades aplicativas de carácter integrador	70
Trabajo autónomo (Lectura de temas/recursos y estudio, resolución de casos y proyecto)	100
Tutorías, seguimiento y evaluación de actividades	50

## 7. EVALUACIÓN

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

Sistema de evaluación	Peso
SE1-Exámenes y tests. Se utilizarán para la evaluación del conocimiento declarativo. Podrán ser pruebas de conocimientos presenciales o incluso propuestos y entregados por medios electrónicos, tanto en forma oral como escrita	30%

SE2-Elaboración de artículos, informes o memorias de diseños. Se utilizarán para la evaluación del conocimiento declarativo y procedimental.	30%
SE3-Evaluación entre compañeros de las simulaciones y de los casos prácticos..	30%
SE4-Las experiencias de campo, conferencias, visitas a empresas e instituciones se evaluarán sobre las bases de las intervenciones en un foro de discusión	5%
SE6-Para la evaluación de las competencias básicas y generales correspondientes a la materia, se utilizarán rúbricas o tests, que podrán ser aplicados por el profesor o mediante sistemas de evaluación entre compañeros	5%

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:

#### Modalidad presencial:

Actividades evaluables	Fecha
A1. Análisis de datos	Semana 4
A2. Revisión de ejemplos de Estructuras de Datos	Semana 6
A3. Análisis descriptivo de datos en Python	Semana 8

A4. Examen	Final Trimestre
A5. Expresividad de modelos de aprendizaje	Semana 2
A6. Aprendizaje con Python	Semana 4
A7. PySpark con Jupyter notebooks	Semana 6
A8. Evaluación del aprendizaje supervisado	Semana 8
A9. Revisión de técnicas de ingeniería de variables	Semana 10
A10. Práctica colaborativa	Semana 12
Examen Final	Al finalizar el tercer trimestre

#### Modalidad online:

Actividades evaluables	Fecha
A1. Análisis exploratorio	Final Unidad Aprendizaje 1
A2. Revisión de ejemplos de Estructuras de Datos	Final Unidad Aprendizaje 2
A3. Análisis descriptivo de datos en Python	Final Unidad Aprendizaje 3
A4. Revisión de técnicas de detección y resolución de problemas de calidad de dato Principales conceptos de caso de negocio	Final Unidad Aprendizaje 4
A5. Revisión del análisis de datos textuales	Final Unidad Aprendizaje 5
A6. Elaboración de un proyecto de análisis y limpieza del dato	Final Unidad Aprendizaje 6
A7. Tipos de aprendizaje en la práctica	Final Unidad Aprendizaje 7
A8. Técnicas de cálculo de la evaluación del aprendizaje supervisado	Final Unidad Aprendizaje 8
A9. Revisión de técnicas de ingeniería de variables	Final Unidad Aprendizaje 8
A10. Revisión de técnicas de aprendizaje supervisado	Final Unidad Aprendizaje 10
A11. Revisión de otras técnicas de aprendizaje	Final Unidad Aprendizaje 11
Examen Final	Al finalizar el tercer trimestre

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

## 9. BIBLIOGRAFÍA

- WES MCKINNEY (2016). Python for Data Analysis. O'Reilly

- JAKE VANDERPLAS (2016). Python Data Science Handbook. O'Reilly
- WILLI RICHERT (2013). Building Machine Learning Systems With Python. Packt Publishing.
- WHITE, T. (2015). Hadoop: The Definitive Guide. O'Reilly.
- ODESKY, M., SPOON, L. & VENNERS, B. (2008). Programming in Scala. Ed. Artima.
- ODESKY, M. (2010). Scala by Example. Programming Method Laboratory, Switzerland: EPFL.
- WAMPLER, D. & PAYNE, A. (2009). Programming Scala. Ed. O'Reilly.
- KARAU, H., KONWINSKI, A., WENDELL, P. & ZAHARIA, M. (2015). Learning Spark. O'Reilly.
- BIHANIC, D. (2015). New Challenges for Data Design. Springer.
- BIHANIC, D. (2015). Empowering Users through Design. Springer.
- FISHER, D., MEYER, M. (2016). Making sense of data. O'Reilly.
- GRAY, J., BOUNEGRU, L., CHAMBERS, L. (2012). The Data Journalism Handbook. O'Reilly.
- ILIINSKY, N., STEELE, J. (2011). Designing Data Visualizations. O'Reilly.
- MUNZNER, T. (2014). Visualization Analysis & Design. CRC Press.

## 10. UNIDAD DE ATENCIÓN A LA DIVERSIDAD

Estudiantes con necesidades específicas de apoyo educativo:

Las adaptaciones o ajustes curriculares para estudiantes con necesidades específicas de apoyo educativo, a fin de garantizar la equidad de oportunidades, serán pautadas por la Unidad de Atención a la Diversidad (UAD).

Será requisito imprescindible la emisión de un informe de adaptaciones/ajustes curriculares por parte de dicha Unidad, por lo que los estudiantes con necesidades específicas de apoyo educativo deberán contactar a través de: [unidad.diversidad@universidadeuropea.es](mailto:unidad.diversidad@universidadeuropea.es) al comienzo de cada semestre.

## 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.