

DATOS GENERALES DE LA ASIGNATURA

ASIGNATURA: Introducción al Análisis de Grandes Volúmenes de Datos, Big Data

PLAN DE ESTUDIOS: Grado en Ingeniería Informática (PGR-INFORM)

GRUPO: 2526-M1

CENTRO: Escuela Politécnica Superior

CARÁCTER DE LA ASIGNATURA: Optativo

ECTS: 6,0

CURSO: 3º

SEMESTRE: 1º Semestre

IDIOMA EN QUE SE IMPARTE:

Castellano, excepto las asignaturas de idiomas que se impartirán en el idioma correspondiente

DATOS DEL PROFESOR

NOMBRE Y APELLIDOS: INMACULADA FIERRO LORENZO

EMAIL: ifierro@uemc.es

TELÉFONO: 983 00 10 00

HORARIO DE TUTORÍAS: Jueves a las 13:00 horas

CV DOCENTE:

Profesora en la Universidad Europea Miguel de Cervantes (UEMC) desde el curso 2016-17.

Acreditación en las figuras de Profesor Contratado Doctor/ Profesor de Universidad Privada por ACSUCYL y ANECA.

CV PROFESIONAL:

.

CV INVESTIGACIÓN:

.

DATOS ESPECÍFICOS DE LA ASIGNATURA

DESCRIPTOR DE LA ASIGNATURA:

Esta asignatura optativa de 6 ETCS forma parte del módulo Complementos de Formación y está ubicada dentro de la Materia 13, Aprendizaje automático y análisis de grandes volúmenes de datos, Big Data, de la memoria del Grado en Ingeniería Informática.

En la asignatura se realizará una introducción a Big Data y las herramientas y lenguajes más utilizados hoy en día para almacenar y analizar la gran cantidad de datos a nuestro alcance y que no se pueden llevar a cabo, en un tiempo prudencial, con los sistemas tradicionales.

No es necesario que el alumno tenga conocimientos previos, pero es aconsejable un nivel básico de estadística, ya que se introducirán distintos tipos de análisis estadísticos multivariantes.

CONTENIDOS DE LA ASIGNATURA:

1. **Bloque I: Primera parte de la asignatura**
 1. Fundamentos en el análisis masivo de datos.
 2. Introducción al análisis multivariante.

OBSERVACIONES CONTENIDO DE LA ASIGNATURA:

El primer bloque de la asignatura permite trabajar los contenidos asociados a las dos primeras partes de los contenidos reflejados en la memoria del grado:

- Fundamentos en el análisis masivo de datos.
- Introducción al análisis multivariante.

Cada tema teorico se complementará con una serie de prácticas, planteadas como casos de estudio, que permitirán al alumno aplicar y asentar los conocimientos adquiridos.

RECURSOS DE APRENDIZAJE:

- Presentaciones de la asignatura proporcionados por el profesor y bibliografía recomendada.
- Laboratorio informático: en él se impartirán todas las clases, cada alumno dispondrá de un ordenador con el que llevar a cabo el seguimiento de las explicaciones sobre los temas objeto de estudio. El profesor utilizará el cañón y la pizarra para exponer los temas.
- Software específico de cálculo estadístico: R/R-Commander, SPSS
- Software auxiliar: Microsoft Word, Microsoft Excel, LandSchool
- Plataforma Moodle: plataforma donde se colgarán los ejercicios, materiales, enunciados de prácticas, se activarán las entregas de prácticas y se realizarán las pruebas escritas.
- Outlook: gestor de correo proporcionado por la universidad a través de su página web. El alumno recibirá notificaciones y respuestas a sus consultas online por esta vía.
- Teams: herramienta que se utilizará en las sesiones virtuales de tutorías individuales y grupales. El alumno también podrá recibir notificaciones y respuestas a sus consultas online por esta vía.

COMPETENCIAS Y RESULTADOS DE APRENDIZAJE A ADQUIRIR POR EL ALUMNO

COMPETENCIAS BÁSICAS:

- CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
- CB3. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética

COMPETENCIAS GENERALES:

- CG02. Capacidad y habilidad para la toma de decisiones en el ámbito tecnológico
- CG03. Capacidad para trabajar en equipos en el ámbito tecnológico.

COMPETENCIAS ESPECÍFICAS:

- COM7. 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.
- CEN2. Capacidad para conocer y aplicar técnicas de análisis estadístico multivariante.
- CEN3. Capacidad para manejar programas y herramientas para el análisis estadístico de datos

RESULTADOS DE APRENDIZAJE:

El alumno será capaz de:

- Conocer los fundamentos del análisis de datos multivariante
- Aplicar distintas técnicas de programación para el análisis de datos
- Utilizar tecnologías para el almacenamiento y procesado masivo de datos

BIBLIOGRAFÍA Y RECURSOS DE REFERENCIA GENERALES

BIBLIOGRAFÍA BÁSICA:

- Huck SW. (2007): Reading Statistics and Research. Pearson. ISBN: 978-02-0551-067-2
- Milton, J. S. y Arnold, J. C. (2003): Probabilidad y estadística con aplicaciones para ingeniería y ciencias computacionales. . Mc Graw Hill. ISBN: 9701043081
- Caballero, R. y Martín, E. (2015): Las bases del Big Data. Catarata. ISBN: 978-8490970867

WEBS DE REFERENCIA:

Web / Descripción

[National Center for Biotechnology Information](https://www.ncbi.nlm.nih.gov). (https://www.ncbi.nlm.nih.gov)
 instituto de investigación, base de datos biológica y repositorio de datos.

[Cran R-Project](https://cran.r-project.org/)(https://cran.r-project.org/)
 Instalación del programa R.

[Métodos Estadísticos con R y R Commander](https://cran.r-project.org/doc/contrib/saez-castillorrcmdrv21.pdf). (https://cran.r-project.org/doc/contrib/saez-castillorrcmdrv21.pdf)
 Métodos Estadísticos con R y R Commander: Texto descargable en pdf.

OTRAS FUENTES DE REFERENCIA:

Competiciones de minería de datos: <http://www.kaggle.com>

PLANIFICACIÓN DEL PROCESO DE ENSEÑANZA-APRENDIZAJE DE LA ASIGNATURA

METODOLOGÍAS:

MÉTODO DIDÁCTICO:

La clase magistral será la actividad principal utilizada para transmitir conocimientos teóricos al principio de cada uno de los temas de la asignatura. Tendrá lugar en el laboratorio de informática asignado a la asignatura.

MÉTODO DIALÉCTICO:

Las actividades prácticas de aprendizaje se utilizarán para asentar los conocimientos explicados. En ellas, el profesor entregará un supuesto práctico y el alumno, además de analizar los datos utilizando un programa estadístico, deber sacar conclusiones aplicando los conocimientos adquiridos en la lección magistral. Se fomentará la participación y el dialogo de los alumnos en la propuesta del análisis de los datos y en la puesta en común de los resultados obtenidos.

MÉTODO HEURÍSTICO:

Los alumnos manejarán bases de datos en el laboratorio de informática y se fomentará la iniciativa del alumno en la resolución de los problemas así como en la elección del programa y tipo de análisis más apropiado para realizar las tareas propuestas.

Como parte de la asignatura se programarán una serie de prácticas (casos de estudio) en las que se aplicarán de manera práctica los conocimientos adquiridos en la asignatura.

CONSIDERACIONES DE LA PLANIFICACIÓN:

Notas previas

La docencia y la evaluación en la asignatura se desarrollarán de forma presencial.

Las fechas de los exámenes finales, tanto en la convocatoria ordinaria como en la extraordinaria no son modificables.

Las clases tendrán una duración de 50 minutos comenzando a las horas "en punto".

El teléfono móvil no podrá utilizarse en el aula excepto cuando exista una indicación explícita del profesor por ser necesario para la realización de alguna actividad didáctica (como la realización de Kahoots).

La planificación estimada, presentada a continuación, podrá verse modificada por causas ajenas a la organización académica. El profesor informará convenientemente a los alumnos de las modificaciones puntuales. Las fechas de

los exámenes finales, tanto en convocatoria de junio (ordinaria) como en convocatoria de julio (extraordinaria) no son modificables.

La primera parte de la asignatura constituirá un bloque (Bloque I) dividido en los temas:

- Fundamentos en el análisis masivo de datos (semanas 1 y 2).
- Introducción al análisis multivariante (semanas 3 a 7).

La evaluación de este bloque se realizará mediante la entrega de dos prácticas, respectivamente, en las semanas 3 y 6 y la realización de una prueba escrita en la semana 7.

A lo largo del curso se pondrán a disposición de los alumnos, en el servicio de reprografía y en la plataforma Moodle, las presentaciones teóricas utilizadas en las clases expositivas para facilitar el seguimiento de la asignatura. Las clases expositivas tendrán una duración de 50 minutos, comenzándose a la hora en punto.

El enunciado de los ejercicios y problemas propuestos en las clases prácticas también serán “colgados” en la plataforma Moodle, así como todo aquello que el docente considere oportuno para el adecuado seguimiento de la asignatura, como pueden ser preguntas de autoevaluación o foro de dudas.

Las tutorías individuales, fijado un horario en la guía docente del profesor y previa solicitud por parte de los alumnos, se podrán realizar de forma presencial en la UEMC o a través de una sesión de Teams, a criterio del profesorado.

Las tutorías académicas grupales están fijadas en la semana amarilla de preparación para la convocatoria ordinaria (2 horas) y extraordinaria (2 horas). Desde la Escuela Politécnica Superior se notificarán, tanto al profesorado como al alumnado, los calendarios de estas tutorías como viene siendo habitual.

PROGRAMACIÓN DE ACTIVIDADES Y EVALUACIONES:

PROGRAMACIÓN DE ACTIVIDADES:

Actividad	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	¿Se evalúa?	CO	CE
Práctica 1			X													X	X	X
Práctica 2						X										X	X	X
Prueba escrita							X									X	X	X

CONSIDERACIONES DE LA EVALUACIÓN EN LA CONVOCATORIA ORDINARIA:

La asignatura consta de dos partes. Para superar la asignatura será necesario obtener al menos un 5 sobre 10 en cada una de las partes. En caso de no superar el 5.0 en alguna de las partes, el alumno no superará la asignatura y la calificación final máxima en la asignatura será de un 4 con independencia de las calificaciones obtenidas. La calificación se obtendrá realizando la media ponderada según los porcentajes especificados en esta guía pero si la calificación supera el 4, la calificación final será de un 4.

Primera parte de la asignatura:

Durante el curso se llevarán a cabo una evaluación continua (2 prácticas) que será necesario superar, además de la prueba escrita que se realizará al final de la primera mitad de la asignatura, para eliminar la parte correspondiente a esta materia, antes de la prueba final de convocatoria ordinaria. En caso de no superar la evaluación continua, en la convocatoria ordinaria de febrero podrá recuperar esta parte de la asignatura.

Prácticas: Las pruebas prácticas podrán ser realizadas en grupos previamente establecidos o individualmente, pero cada alumno las subirá individualmente a la plataforma Moodle. Los ejercicios deberán ser realizados o subidos a la plataforma antes de la fecha límite indicada. En caso de haber algún problema, el alumno enviará un correo electrónico comunicando la incidencia a su profesora, siempre antes de la fecha de cierre de la tarea. No se admitirán entregas a través del correo electrónico. Cada práctica sin entrega será calificada con 0 puntos. Los ejercicios propuestos en las prácticas serán resueltos y discutidos en el aula una vez finalizado el plazo de entrega. La profesora no subirá archivos adicionales con los ejercicios que ya hayan sido resueltos en clase.

Prueba escrita: Se realizará una prueba de evaluación escrita que tendrá carácter eliminatorio, pero siempre que el alumno alcance o supere la nota de 5.0 puntos sobre 10 tanto en esta prueba como en la nota media de las dos prácticas.

La nota media de las prácticas 1 y 2 contribuirá con un 25% a la nota final de la asignatura y la prueba escrita con otro 25 %.

Segunda parte de la asignatura:

La evaluación continua de la segunda parte de la asignatura se llevará a cabo a partir de los siguientes ítems:

- Prácticas: 35%
- Prueba escrita de prácticas: 15%

Se llevarán a cabo una serie de pequeñas prácticas de aplicación de los contenidos de la asignatura que se irán entregando a lo largo del curso. Además, para evaluar los conocimientos adquiridos, se realizará una prueba escrita de prácticas. Será necesario obtener al menos un 4 sobre 10 en esta prueba para poder superar la asignatura. A efectos de cálculo de la calificación final, la prueba no realizada o las prácticas no entregadas o entregadas fuera de plazo contabilizarán con una calificación de 0.

Convocatoria ordinaria:

Los alumnos que no hayan superado la primera parte de la asignatura realizarán una parte práctica y una prueba escrita en la convocatoria ordinaria de febrero, tanto si superaron las prácticas o la prueba escrita, como si suspendieron ambas partes. Para superar la asignatura deberán obtener una nota de, al menos, 5.0 puntos sobre 10 que supondrá el 50% de la ota final de la asignatura.

Los alumnos que no hayan superado la segunda parte de la asignatura (independientemente de las notas que hayan obtenido tanto en la prueba como en las prácticas), deberán presentarse a una prueba de ejecución práctica en la fecha oficial de convocatoria ordinaria de la asignatura que deberán superar (obtener al menos un 5 sobre 10) y que supondrá el 50% de la nota final de la asignatura.

La planificación de la evaluación tiene un carácter meramente orientativo y podrá ser modificada a criterio del profesor, en función de circunstancias externas y de la evolución del grupo. Los sistemas de evaluación descritos en esta GD son sensibles tanto a la evaluación de las competencias como de los contenidos de la asignatura.

La realización fraudulenta de cualquiera de las pruebas de evaluación, así como la extracción de información de las pruebas de evaluación, será sancionada según lo descrito en el Reglamento 7/2015, de 20 de noviembre, de Régimen Disciplinario de los estudiantes, Arts. 4, 5 y 7 y derivarán en la pérdida de la convocatoria correspondiente, así como en el reflejo de la falta y de su motivo en el expediente académico del alumno.

CONSIDERACIONES DE LA EVALUACIÓN EN LA CONVOCATORIA EXTRAORDINARIA:

En caso de no aprobar en febrero, la prueba evaluación de la convocatoria extraordinaria del mes de julio consistirá en una prueba escrita de cada una de las partes de la asignatura no superadas en convocatoria ordinaria.

Al igual que en la Convocatoria Ordinaria, en la Convocatoria Extraordinaria, la planificación de la evaluación tiene un carácter meramente orientativo y podrá ser modificada a criterio del profesor, en función de circunstancias externas y de la evolución del grupo.

Los sistemas de evaluación descritos en esta Guía Docente son sensibles tanto a la evaluación de las competencias como de los contenidos de la asignatura.

La realización fraudulenta de cualquiera de las pruebas de evaluación, así como la extracción de información de las pruebas de evaluación, será sancionada según lo descrito en el Reglamento 7/2015, de 20 de noviembre, de Régimen Disciplinario de los estudiantes, Arts. 4, 5 y 7 y derivarán en la pérdida de la convocatoria correspondiente, así como en el reflejo de la falta y de su motivo en el expediente académico del alumno.

SISTEMAS DE EVALUACIÓN:

SISTEMA DE EVALUACIÓN	PORCENTAJE (%)
Ejecución de prácticas	50%
Pruebas escritas	50%