

24-25

GUÍA DE ESTUDIO PÚBLICA



INTRODUCCIÓN A LA CIENCIA Y EL ANÁLISIS DE DATOS

CÓDIGO 21580042

UNED

24-25**INTRODUCCIÓN A LA CIENCIA Y EL
ANÁLISIS DE DATOS****CÓDIGO 21580042**

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN
REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA
ASIGNATURA
EQUIPO DOCENTE
HORARIO DE ATENCIÓN AL ESTUDIANTE
COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE
RESULTADOS DE APRENDIZAJE
CONTENIDOS
METODOLOGÍA
SISTEMA DE EVALUACIÓN
BIBLIOGRAFÍA BÁSICA
BIBLIOGRAFÍA COMPLEMENTARIA
RECURSOS DE APOYO Y WEBGRAFÍA
IGUALDAD DE GÉNERO

Nombre de la asignatura	INTRODUCCIÓN A LA CIENCIA Y EL ANÁLISIS DE DATOS
Código	21580042
Curso académico	2024/2025
Título en que se imparte	MÁSTER UNIVERSITARIO EN FÍSICA AVANZADA
Tipo	CONTENIDOS
Nº ETCS	6
Horas	150
Periodo	SEMESTRE 1
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

La asignatura “Introducción a la Ciencia y el Análisis de Datos” aborda a un nivel introductorio algunas de las técnicas más populares que se utilizan actualmente en el tratamiento de datos para obtener de ellos información relevante.

Se estudiarán las algunas de las técnicas de utilidad más general, como la estadística descriptiva, las regresiones lineal, polinomial, multinomial y logística, el análisis de correspondencias, los árboles de decisión, la clusterización y modelos de aprendizaje estadístico. Se aprenderá a programar y utilizar estos métodos sobre conjuntos de datos y/o para generar modelos y obtener de ellos información relevante sobre sistemas de interés.

La asignatura es de utilidad para todos aquellos estudiantes que tengan interés en el análisis de datos, en el tratamiento computacional de la información o en la construcción de modelos. Aunque estas técnicas son de enorme utilidad en Física, Ingeniería y otras disciplinas científicas su carácter es general y son también empleados en áreas muy dispares como la Economía o la Sociología y realmente en cualquier otra en la que existan conjuntos de datos numéricos susceptibles de ser analizados para recabar información.

La asignatura es optativa, impartándose en el primer cuatrimestre del Máster, y consta de 6 ECTS, equivalentes a 150 horas de trabajo. El enfoque de la asignatura es fundamentalmente práctico, de manera que, a título orientativo, dichas horas de trabajo se distribuyen de la siguiente manera:

- Trabajo autónomo de los contenidos teóricos (lectura y consulta de los materiales didácticos; estudio crítico de los mismos; realización de los ejercicios de autoevaluación): 50%
- Realización de las actividades prácticas y elaboración de los informes de resultados: 50%.

Dentro del presente Máster, esta asignatura proporciona conocimientos y herramientas matemáticas y computacionales para el tratamiento y representación de datos, tanto experimentales como resultados de simulación, y disponibles en diferentes bases de datos.

Como ya se ha comentado, estas técnicas son de aplicación general y no se restringen a los

diferentes campos de la Física, pero dentro de esta área cabe destacar que han sido usadas en el análisis de los datos de Física de Partículas en el CERN, en Sociofísica y en muchos otros campos.

La asignatura pertenece al Módulo “Física Computacional” pero sus contenidos y aplicaciones son de carácter tan general que puede resultar de enorme utilidad para otros itinerarios ya que faculta al estudiante con habilidades de análisis de datos aplicables a una infinidad de problemas diferentes.

Toda la información contenida en el CURSO VIRTUAL, específicamente modalidad y formato de entrega de las tareas o PECs, realización de las pruebas presenciales, fechas de entrega, etc. es de obligado conocimiento por parte de los estudiantes de la asignatura. Esto incluye la información dada por el Equipo Docente en los diferentes foros con los que se comunica con el estudiante. Toda la normativa contenida en el CURSO VIRTUAL es de obligado cumplimiento para todos los estudiantes de la asignatura.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Para abordar la asignatura con garantías de éxito son precisos conocimientos avanzados en Física y Matemáticas, que hayan sido adquiridos en asignaturas de grados o licenciaturas en Ciencias o Ingeniería. En particular:

- 1.- Álgebra lineal (al nivel de estudios de algunos grados en ingeniería o ciencias).
- 2.- Análisis matemático (al nivel de estudios de algunos grados en ingeniería o ciencias).
- 3.- Probabilidad y estadística básicas (al nivel de estudios de algunos grados en ingeniería o ciencias).
- 4.- Cálculo numérico y programación (al nivel de estudios de algunos grados en ingeniería o ciencias).

En general, los conocimientos adquiridos en grados o licenciaturas en Ciencias Físicas o Químicas, Matemáticas y las Ingenierías pudieran ser suficientes.

Por otra parte, el estudiante ha de estar bien familiarizado con el uso de ordenadores, ya que buena parte del trabajo de la asignatura (y de las tareas que permiten la evaluación de la misma) está basado en la ejecución y modificación de programas de cálculo. El software de uso en las tareas de la asignatura será Matlab/Octave, este software es el más adecuado para plasmar los conceptos físicos y matemáticos en forma de programas y algoritmos. La UNED proporciona pleno acceso a Matlab que incluye soporte técnico para los estudiantes. En la asignatura, además, se proveerá de material introductorio en Octave/Matlab que capacite para la programación de los algoritmos básicos más usados. Es muy conveniente tener conocimientos generales de programación científica para trabajar más cómodamente en la asignatura.

En la asignatura se usará bibliografía que se encuentra exclusivamente en idioma inglés.

EQUIPO DOCENTE

Nombre y Apellidos	DAVID GARCIA ALDEA (Coordinador de asignatura)
Correo Electrónico	dgaldea@fisfun.uned.es
Teléfono	91398-7636
Facultad	FACULTAD DE CIENCIAS
Departamento	FÍSICA FUNDAMENTAL

HORARIO DE ATENCIÓN AL ESTUDIANTE

El medio básico de comunicación y tutorización entre estudiantes y equipo docente son las herramientas de comunicación del Curso virtual, especialmente los Foros de debate.

Además, podrán utilizarse el correo electrónico, reunión virtual (TEAMS) y la entrevista personal.

Nota importante: el equipo docente puede cambiar con posterioridad a la redacción de esta información. En todo caso, los profesores que constan en el apartado "Equipo docente" están actualizados.

Profesor: David García Aldea

E-mail: dgaldea@fisfun.uned.es

Teléfono: 91 398 7136

Horario: Martes, de 16 a 20 h

Edificio Biblioteca UNED, planta 1 (Mediateca). Paseo Senda del Rey 5. 28040 Madrid, España

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

COMPETENCIAS

CM1 Poseer la capacidad para el desarrollo de una aptitud crítica ante el aprendizaje que le lleve a plantearse nuevos problemas desde perspectivas no convencionales.

CM2 Adquirir los conocimientos necesarios en Física Avanzada para incorporarse a un grupo de investigación o a empresas.

CM3 Adquirir la capacidad para abordar y resolver un problema avanzado en la física teórica, computacional o de fluidos, mediante la elección adecuada del contexto teórico, la identificación de los conceptos relevantes y el uso de las técnicas matemáticas que constituyen la mejor aproximación para así llegar a la solución.

CM5 Analizar problemas nuevos en sistemas poco conocidos y determinar similitudes y diferencias con modelos de referencia.

CM6 Analizar críticamente resultados experimentales, analíticos y numéricos en el campo de

la física avanzada.

RESULTADOS DE APRENDIZAJE

CONOCIMIENTOS O CONTENIDOS

CN1 Comprender conceptos avanzados de Física y demostrar, en un contexto de investigación científica altamente especializada, una relación detallada y fundamentada entre los aspectos teóricos y prácticos y la metodología empleada en este campo.

CN2 Conocer y comprender los elementos más relevantes de la física teórica, computacional y de fluidos actual. Profundizar en la comprensión de las teorías que se encuentran en la frontera de estos temas, incluyendo su estructura matemática, su confrontación con resultados experimentales, y la descripción de los fenómenos físicos que dichas teorías explican.

CN3 Conocer los sistemas operativos y lenguajes de programación y herramientas de computación relevantes en el campo de la física avanzada.

HABILIDADES O DESTREZAS

H1 Elaborar un trabajo escrito con datos bibliográficos, teóricos y/o experimentales, escribiendo un resumen o articulado en extenso (tal y como se realizan los artículos científicos), formulando hipótesis razonables, composiciones originales y conclusiones motivadas.

H2 Comunicar con claridad y rigor los resultados de un trabajo de investigación de forma oral o escrita.

H3 Utilizar bibliografía y fuentes de información especializada, propias del ámbito de conocimiento de la física, manejando las principales bases de datos de recursos científicos.

H4 Saber trabajar en equipo y comunicarse con la comunidad académica en su conjunto y con la sociedad en general acerca de la Física Avanzada, tanto en sus implicaciones académicas, productivas o sociales.

H5 Modelizar sistemas de alto grado de complejidad. Identificar variables y parámetros relevantes y realizar aproximaciones que simplifiquen el problema. Construir modelos físicos que describan y expliquen situaciones en ámbitos diversos.

H7 Resolver problemas algebraicos, de resolución de ecuaciones y de optimización mediante métodos numéricos

H8 Modelar y simular fenómenos físicos complejos por ordenador.

COMPETENCIAS

CM1 Poseer la capacidad para el desarrollo de una aptitud crítica ante el aprendizaje que le lleve a plantearse nuevos problemas desde perspectivas no convencionales.

CM2 Adquirir los conocimientos necesarios en Física Avanzada para incorporarse a un grupo de investigación o a empresas.

CM3 Adquirir la capacidad para abordar y resolver un problema avanzado en la física teórica, computacional o de fluidos, mediante la elección adecuada del contexto teórico, la identificación de los conceptos relevantes y el uso de las técnicas matemáticas que

constituyen la mejor aproximación para así llegar a la solución.

CM5 Analizar problemas nuevos en sistemas poco conocidos y determinar similitudes y diferencias con modelos de referencia.

CM6 Analizar críticamente resultados experimentales, analíticos y numéricos en el campo de la física avanzada.

CONTENIDOS

Tema 1: Introducción a la Ciencia de Datos.

Tema 2: Estadística para la Construcción de Modelos y su Evaluación.

Tema 3: Regresión Lineal y Polinomial.

Tema 4: Regresión Logística.

Tema 5: Agrupamiento por Medias.

Tema 6: Reglas de Asociación.

Tema 7: Clasificador Bayesiano.

Tema 8: Árboles de Decisión.

Tema 9: Análisis de Series Temporales.

Tema 10: Modelos de Aprendizaje Estadístico.

Tema 11: Construcción y Testeo de Modelos Predictivos.

METODOLOGÍA

La metodología de la asignatura está basada en la enseñanza a distancia, utilizando el curso virtual implementado en la plataforma docente de la UNED. Dentro de ese curso virtual, los estudiantes dispondrán de:

- 1.- La información general de la asignatura, donde se establece el orden temporal de actividades y prácticas.
- 2.- Material didáctico específico (teórico y práctico) de la asignatura.
- 3.- Enlaces a los recursos informáticos necesarios para la realización de las Tareas prácticas, así como la explicación de lo que se pide en las mismas.
- 4.- Enlaces a material bibliográfico complementario.
- 5.- Herramientas de comunicación: foros de consulta y debate, y plataforma de entrega de los informes de las Tareas prácticas.

Siguiendo el esquema temporal de la asignatura, el estudiante abordará el estudio autónomo de los contenidos teóricos de cada uno de los temas principalmente a partir del texto base de la asignatura y el material complementario del curso virtual pero también de los recursos externos como artículos, fuentes de datos, etc.

El curso se completa con la realización a lo largo del mismo de varias pruebas tipo Test y Tareas prácticas, en las que se usan herramientas informáticas, tanto en la realización de los cálculos como en la escritura de las memorias. En los trabajos prácticos se aplicarán a técnicas matemáticas y los conocimientos teóricos adquiridos a la construcción de modelos, la extracción de información de relevancia de conjuntos de datos y la resolución de problemas.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen

No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad

No

Descripción

No hay examen final de la asignatura.

Criterios de evaluación

Ponderación de la prueba presencial y/o los trabajos en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC?

Si, PEC presencial

Descripción

Existen tres tipos de tareas en esta asignatura

Los Test Online consisten en pruebas objetivas de varias preguntas de respuesta múltiple. En ellos se pondrá a prueba el conocimiento de los estudiantes sobre los conceptos teóricos fundamentales y su aplicación. Pueden versar sobre la materia del texto base o los artículos.

Los Trabajos Prácticos consisten en la resolución por parte del estudiante de ejercicios o problemas en los cuales será necesario usar un software de programación -Matlab u Octave- para obtener los resultados. Debe finalmente redactarse una pequeña memoria en la que figuren los resultados, una discusión pertinente y las conclusiones que se pueden sacar a partir de ellos.

El Trabajo Final de la asignatura consistirá en un Trabajo Práctico al estilo de los ya realizados pero con un enunciado más abierto y cierta libertad por parte del estudiante. En el podrá poner en práctica los conocimientos y las destrezas adquiridas durante el curso de la asignatura.

Se encontrarán distribuidos según los bloques de la asignatura para que cubran todo el temario de esta.

Criterios de evaluación

Test Online (aproximadamente 25%)

Trabajos prácticos (aproximadamente 45%)

Trabajo de Investigación final (aproximadamente 30%)

Ponderación de la PEC en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

El curso está estructurado en trece semanas, siendo la primera semana la que comienza con la inauguración del curso académico y la decimotercera la que termina justo antes de la primera semana de exámenes según el calendario oficial de la UNED. El periodo no lectivo de vacaciones de Navidad suele encontrarse tras la semana 11.

Tests:

Se trata de pruebas sencillas de una hora de duración con preguntas de respuesta múltiple. Se pueden realizar simplemente habiendo comprendido el material, y su objetivo es que los estudiantes vayan siguiendo el curso con el mínimo de conocimientos imprescindible. Incluyen el temario que se ha estudiado hasta la semana anterior y se pueden realizar utilizando cualquier tipo de material que el estudiante desee. Estarán disponibles durante un tiempo estipulado en la semana que les corresponda. Las fechas aproximadas de realización de los tests son:

Test 1 (5%): Semana 3

Test 2 (5%): Semana 5

Test 3 (5%): Semana 7

Test 4 (5%): Semana 9

Test 5 (5%): Semana 11

Tareas:

Las tareas consisten en la realización de un conjunto de problemas académicos, estudio de un problema más extenso o de un artículo y la redacción de una memoria con los resultados que se entregará en el curso virtual. Los problemas a realizar versarán sobre el temario estudiado hasta la fecha y requieren trabajo de programación en Matlab. Las fechas de entrega de las memorias corresponderán al final de las semanas aproximadas que se indican:

Tarea 1 (15%): Semana 6

Tarea 2 (15%): Semana 8

Tarea 3 (15%): Semana 10

Trabajo Final:

El trabajo final de la asignatura consiste en el estudio de un problema o tema específico de interés del estudiante y la realización de una memoria.

En la semana 10 se presentará una propuesta de proyecto final por parte del alumno en el que pueda poner en práctica los conocimientos adquiridos en la asignatura. El equipo docente evaluará la viabilidad y adecuación de la propuesta como proyecto final.

Durante las semanas restantes el estudiante trabajará en el desarrollo del proyecto final y finalmente entregará una memoria con los resultados y las pertinentes discusiones y conclusiones que haya obtenido. La semana aproximada de entrega del proyecto será:

Trabajo final (30%): Semana 13

Para poder aprobar la asignatura es imprescindible haber entregado todas las

tareas y el trabajo final. Es necesario aprobar cada una de las tareas y el trabajo final por separado. Si no se realiza alguno de los tests, la asignatura puede ser superada, pero se pierde esa puntuación. Los tests realizados siempre contribuyen positivamente a la nota final aportando su fracción de la puntuación, aunque se obtenga una calificación inferior a cinco.

Evaluación en la convocatoria extraordinaria:

Los tests de la asignatura deben ser realizados en la fecha en la que figuran en el curso. Su nota se guarda para la convocatoria extraordinaria y no hay posibilidad de repetirlos.

Las notas de las tareas superadas se guardan para la convocatoria extraordinaria y las que estuvieran pendientes se deben entregar con fecha del 1 de julio. Se permite también entregar en la misma fecha las tareas que se hubieran aprobado, pero que el estudiante desee mejorar.

Si no se tuviese ya una propuesta de trabajo final válido se puede presentar hasta el 1 de julio. El equipo docente la evaluará antes del 15 de julio y el estudiante podrá trabajar el proyecto final. Este deberá ser entregado con fecha límite del 10 de septiembre.

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

Criterios de evaluación

Ponderación en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

El sistema de evaluación está basado en el sistema de evaluación continua. Es obligatorio la realización de todos las Tareas Prácticas de la asignatura para poder aprobar. Finalmente se realizará un Trabajo Final de enunciado más abierto en el que se pongan en práctica los conocimientos, destrezas y habilidades desarrolladas durante el curso de la asignatura.

BIBLIOGRAFÍA BÁSICA

ISBN(13): 9781118876138

Título: DATA SCIENCE AND BIG DATA ANALYTICS: DISCOVERING, ANALYZING, VISUALIZING AND PRESENTING DATA 1st Edition. May 17th 2015 edición

Autor/es: Emc Education Services (Redactor)

Editorial: JOHN WILEY & SONS INC

ISBN(13): 9788412619850

Título: EL ARTE DE LA ESTADÍSTICA: CÓMO APRENDER DE LOS DATOS. N.º 1 edición (15 mayo 2023) edición

Autor/es: David Spiegelhalter

Editorial: CAPITAN SWING S.L.

El primer texto base de la asignatura (disponible exclusivamente en inglés) es:

"Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data." EMC Education Services (Editor). 1st Edition. Wiley (2017). ISBN: 978-1-118-87613-8

El primer texto cubre los principales temas de la asignatura que serán complementados con los materiales en el curso virtual.

El segundo texto base de la asignatura está disponible en castellano, aunque también puede utilizarse la edición en inglés.

Para el estudio de la asignatura han de utilizarse ambos textos y el material complementario del curso virtual. El primer texto cubre los aspectos técnicos de la asignatura acerca de la aplicación e implementación computacional de los distintos métodos de Ciencia de Datos.

El segundo texto, en castellano o inglés es:

"El Arte de la Estadística: Aprendiendo de los Datos." David Spiegelhalter. CAPITAN SWING S.L; N.º 1 edición (15 mayo 2023). ISBN: 978-8412619850.

Este segundo texto complementa al primero en las estrategias a seguir al estudiar fuentes de datos, plantear experimentos e interpretación de los resultados.

A lo largo del curso se trabajará la asignatura estudiando cada semana unos capítulos del primer texto y otros del segundo y el material especificado en el curso virtual, que consiste sobre todo en ejercicios resueltos planteados en ambos textos.

BIBLIOGRAFÍA COMPLEMENTARIA

Bibliografía complementaria de la asignatura que puede ser interesante para los estudiantes son las referencias:

- Data Science from Scratch: First Principles with Python 1st Edition. O'Reilly Media (2015).
- The Data Science Handbook. Field Cady. Wiley (2017).
- Python Data Science Essentials - Learn the fundamentals of Data Science with Python. Alberto Boschetti, Luca Massaron. Packt Publishing (2015).

RECURSOS DE APOYO Y WEBGRAFÍA

Todos los recursos de apoyo al estudio están contenidos en la plataforma virtual.

El estudiante ha de prestar particular atención a:

- 1.- Los contenidos teóricos básicos de la asignatura, tanto en aquellos aportados en el curso como en los que están la bibliografía recomendada.
- 2.- Guiones de las Tareas (trabajos prácticos).
- 3.- Enlaces a los artículos que constituyen la bibliografía complementaria.

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.