

23-24

MÁSTER UNIVERSITARIO EN
INGENIERÍA Y CIENCIA DE DATOS

GUÍA DE ESTUDIO PÚBLICA



APRENDIZAJE AUTOMÁTICO I

CÓDIGO 31110037

UNED

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	

Nombre de la asignatura	APRENDIZAJE AUTOMÁTICO I
Código	31110037
Curso académico	2023/2024
Título en que se imparte	MÁSTER UNIVERSITARIO EN INGENIERÍA Y CIENCIA DE DATOS
Tipo	CONTENIDOS
Nº ECTS	6
Horas	150.0
Periodo	SEMESTRE 1
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

PRESENTACIÓN

El aprendizaje automático es clave para el desarrollo de sistemas inteligentes y el análisis de datos en ciencia e ingeniería. Tiene un papel decisivo en tecnologías de uso cada vez más común como los reconocedores de voz de los teléfonos móviles o los vehículos autónomos, por ejemplo, así como también sirve para permitir el acceso a la información contenida en nuestro ADN o para dotar de sentido a la avalancha de datos que pueden recogerse en la web.

Así, dado que el aprendizaje automático es una de las bases de la ciencia de datos, en este curso se ofrece una introducción a los métodos más comunes que son de uso frecuente en este ámbito. En esta asignatura se tratarán los fundamentos teóricos, así como algoritmos esenciales en el aprendizaje supervisado y una introducción al aprendizaje no supervisado. Proporcionará las competencias para resolver problemas de aplicación de estas tecnologías y la base necesaria para abordar el estudio de otras materias del Máster relacionadas que se cursan en el segundo semestre.

En el aprendizaje supervisado se utilizan pares de datos de entrada y de salida con los que se entrenan los distintos algoritmos. Tras el entrenamiento el sistema debe ser capaz de asociar un dato de entrada, no visto previamente, con su dato de salida.

CONTEXTUALIZACIÓN

La asignatura tiene 6 créditos ECTS, de carácter obligatorio, impartida en el primer semestre del Máster Universitario en Ingeniería y Ciencia de Datos. Guarda relación con las siguientes asignaturas también disponibles en el mismo Máster:

- Modelado estadístico de datos
- Programación en entornos de datos
- Aprendizaje automático II
- Deep Learning

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Se recomienda que los interesados en cursar el Máster tengan un nivel de lectura en inglés suficiente como para entender contenidos técnicos en dicha lengua.

Gran parte de la bibliografía, así como los recursos proporcionados al estudiante en el curso virtual pueden estar únicamente en inglés, debido a la novedad de algunos de los contenidos propuestos para la asignatura.

Se fomentará el uso de software libre siempre y cuando sea posible para la realización de las actividades y las practicas propuestas. El lenguaje de programación que se usará será Python y varias herramientas que usan este lenguaje como scikit-learn y Keras. Se proporcionará material complementario con introducción a la programación en Python.

EQUIPO DOCENTE

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

JOSE MANUEL CUADRA TRONCOSO (Coordinador de asignatura)
jmcuadra@dia.uned.es

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

ANTONIO RODRIGUEZ ANAYA
arodriguez@dia.uned.es

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

JORGE PEREZ MARTIN
jperezmartin@dia.uned.es

HORARIO DE ATENCIÓN AL ESTUDIANTE

La tutorización de los estudiantes tendrá lugar esencialmente a través de los foros del curso virtual en Alf, la plataforma virtual de la UNED, aunque también podrán utilizarse ocasionalmente otros medios, tales como chats interactivos, servicios de mensajería instantánea y el correo electrónico. Adicionalmente, está también previsto, para temas personales que no afecten al resto de los estudiantes, atender consultas en persona o por teléfono.

El seguimiento del aprendizaje se realizará revisando la participación de los alumnos en los distintos foros de debate y las aportaciones de material nuevo además de la entrega en fecha de los diferentes trabajos prácticos que se han planificado durante la evolución del curso.

En caso de necesitar contactar con el Equipo Docente por medios distintos al curso virtual, se utilizará preferentemente el correo electrónico, pudiéndose también realizar consultas

telefónicas, videoconferencias y entrevista personal en los horarios establecidos y que se muestran a continuación en la siguiente tabla.

Profesor	Horario de atención	Correo electrónico	Teléfono de contacto	Dirección postal
José Manuel Cuadra Troncoso	Lunes de 16 a 20 horas	jmcuadra@dia.uned.es	91 398 7144	Juan del Rosal, 16, 3-19
Jorge Pérez Martín	Martes de 8 a 12 horas	jperezmartin@dia.uned.es	91 398 9387	Juan del Rosal, 16, 3-01
Antonio Rodríguez Anaya	Lunes de 16 a 20 horas	arodriguez@dia.uned.es	91 398 6550	Juan del Rosal, 16, 3-04

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

COMPETENCIAS BÁSICAS

CB6 - 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.

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB10 - 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.

COMPETENCIAS GENERALES

CG1 - Identificar los métodos apropiados para la solución de problemas asociados a la ciencia de datos y la analítica de información

CG2 - Ser capaz de aplicar diferentes técnicas de aprendizaje máquina, seleccionando el algoritmo óptimo que genere modelos precisos y permita el desarrollo de soluciones predictivas en diferentes ámbitos de uso

CG5 - Utilizar las habilidades de científico de datos y/o ingeniero de datos en entornos de trabajo multidisciplinares y ser capaz de distinguir/organizar las diferentes actividades de los roles en dicho entorno

COMPETENCIAS TRANSVERSALES

CT1 - Ser capaz de abordar y desarrollar proyectos innovadores en entornos científicos, tecnológicos y multidisciplinares.

CT2 - Ser capaz de tomar decisiones y formular juicios basados en criterios objetivos (datos

experimentales, científicos o de simulación disponibles).

COMPETENCIAS ESPECÍFICAS

CE2 - Desarrollar aplicaciones/servicios/scripts orientados a la analítica de datos y analizar el uso de diferentes librerías para el desarrollo e implementación de métodos numéricos, algoritmos y modelos asociados a los datos

CE5 - Desarrollar modelos de aprendizaje máquina (Machine Learning) basados en las diferentes categorías de clasificación: supervisada, no supervisada y semi-supervisada

CE6 - Diseñar mecanismos de evaluación de modelos de aprendizaje y comprender las métricas usadas para dicha evaluación

RESULTADOS DE APRENDIZAJE

Los resultados más relevantes que se pretenden alcanzar con el estudio de esta asignatura son los siguientes:

- Identificar y aplicar los métodos de suavizado kernel unidimensionales y la regresión local.
- Identificar y aplicar las funciones de base radial y los modelos de mezclas para la resolución de problemas de regresión.
- Identificar, construir y ajustar árboles generalizados aditivos para regresión y clasificación.
- Identificar y aplicar la familia de métodos basados en prototipos y sus ventajas e inconvenientes.
- Distinguir los fundamentos teóricos de las redes neuronales artificiales y los conceptos de sobreajuste, inicialización y topología.
- Identificar y aplicar modelos basados en redes neuronales artificiales a problemas reales.
- Distinguir la fundamentación teórica del clasificador de vectores soporte y la clasificación de margen máximo.
- Identificar, construir y ajustar máquinas de vectores soporte para clasificación y regresión.
- Identificar y aplicar las distintas familias de métodos de selección de variables, así como los métodos de reducción de dimensionalidad.

CONTENIDOS

Bloque 1: Métodos básicos de aprendizaje automático

En este bloque se dará una introducción al aprendizaje automático en su tres vertientes: supervisado, no supervisado y semi-supervisado. Se tratarán con cierto detalle algunos algoritmos básicos de aprendizaje no supervisado y supervisado y su implementación con scikit-learn y otras herramientas.

En el curso virtual se dispone de un vídeo de introducción al módulo.

Tema 1 Introducción al Aprendizaje Automático

- Introducción

- Tipos de sistemas de Aprendizaje Automático
- Principales desafíos del Aprendizaje Automático
- Testeo y validación

Tema 2 Introducción a las Herramientas a utilizar

- Instalación de las principales herramientas.
- Introducción a Numpy
- Introducción a Pandas
- Introducción a Matplotlib
- Introducción a Scikit-learn
- Introducción a la metodología de los proyectos de Aprendizaje Automático

Tema 3 Métodos simples de Aprendizaje Automático Supervisado

- Naive Bayes
- k-Nearest-Neighbors
- Árboles de decisión

Tema 4 Métodos simples de Aprendizaje Automático No-Supervisado y Semi-supervisado

- Técnicas de agrupamiento: K-means y DBSCAN
- Modelos de mezclas gaussianas
- Kernel density
- Propagación y extensión de etiquetas

Los contenidos del módulo son introductorios y no tienen una dificultad elevada, siendo sencillos su estudio.

Bloque 2: Selección de variables y reducción de la dimensionalidad

En este bloque se dará una introducción a la selección de variables y reducción de la dimensionalidad, las principales técnicas para su implementación y ejemplos de uso con scikit-learn.

En el curso virtual se dispone de un vídeo de introducción al módulo.

Tema 5 Selección de variables y reducción de dimensionalidad

- Ingeniería de características (variables)
- Introducción a la reducción de la dimensionalidad
- Análisis de componentes principales
- Otras técnicas de reducción de dimensionalidad

Los contenidos del tema tienen una dificultad media ya que necesitan de conocimientos matemáticos para su perfecta comprensión, estos conocimientos se imparten en las titulaciones de acceso al Máster.

Bloque 3: Redes neuronales

En este bloque se dará una introducción a la teoría de las redes neuronales, su implementación usando Keras sobre TensorFlow, aspectos prácticos del uso de redes neuronales y ejemplos de clasificación y regresión.

En el curso virtual se dispone de un vídeo de introducción al módulo.

Tema 6 Introducción a las redes neuronales

- De las neuronas naturales a las artificiales
- El perceptrón
- El perceptrón multicapa y la retropropagación del gradiente
- Regresión y clasificación con perceptrones multicapa

Tema 7 Aspectos prácticos del uso de perceptrones multicapa

- Instalación de las herramientas a utilizar
- Introducción a Tensonflow y Keras
- Ejemplos de regresión y clasificación con Keras
- Uso avanzado de Keras: creación de subclases y uso de callbacks
- Uso de Tensorboard para visualización
- Ajuste fino de hiperparámetros

Los contenidos del tema tienen una dificultad media ya que necesitan de conocimientos matemáticos para su perfecta comprensión, estos conocimientos se imparten en las titulaciones de acceso al Máster.

Bloque 4: Máquinas de vectores soporte

En este bloque se dará una introducción a la teoría de las máquinas de vectores soporte, su implementación usando scikit-learn y su uso para claisificación lineal, no-lineal y regresión.

En el curso virtual se dispone de un vídeo de introducción al módulo.

Tema 8 Máquinas de vectores soporte

- Clasificación lineal
- Clasificación no-lineal
- Regresión
- Fundamentos matemáticos

Los contenidos del tema tienen una dificultad media ya que necesitan de conocimientos matemáticos para su perfecta comprensión, estos conocimientos se imparten en las titulaciones de acceso al Máster.

METODOLOGÍA

Esta asignatura ha sido diseñada para la enseñanza a distancia. Por tanto, el sistema de enseñanza-aprendizaje estará basado en gran parte en el estudio independiente o autónomo del estudiante. Para ello, el estudiante contará con diversos materiales que permitirán su trabajo autónomo y la Guía de Estudio de la asignatura, que incluye orientaciones para la realización de las actividades prácticas. Asimismo, mediante la plataforma virtual de la UNED existirá un contacto continuo entre el equipo docente y los/as estudiantes, así como una interrelación entre los propios estudiantes a través de los foros, importantísimo en la enseñanza no presencial.

El estudio de esta asignatura se realizará a través de los materiales y enlaces que el Equipo Docente publicará en el curso virtual.

Las actividades formativas para el estudio de la asignatura son las siguientes:

- Estudios de contenidos (80 horas)
- Actividades en la plataforma virtual (10 horas)
- Prácticas informáticas (60 horas)
- Total: 150 horas

Los medios necesarios para el aprendizaje son:

- 1. Materiales teórico-prácticos** preparados por el Equipo Docente para cubrir los conceptos básicos del temario.
- 2. Bibliografía complementaria.** El estudiante puede encontrar en ella información adicional para completar su formación.
- 3. Curso Virtual** de la asignatura, donde el estudiante encontrará:
 - Una **guía de la asignatura** en la que se hace una descripción detallada del plan de trabajo propuesto.
 - Un **calendario** con la distribución temporal de los temas propuesta por el Equipo Docente y con las fechas de entrega de las actividades teórico-prácticas que el estudiante tiene que realizar para su evaluación.
 - Enunciado de las **actividades teórico-prácticas** propuestas y zona donde depositar los entregables asociados a dichas actividades.
 - Los **foros** por medio de los cuales el Equipo Docente aclarará las dudas de carácter general y que se usarán también para comunicar todas aquellas novedades que surjan a lo largo del curso. Éste será el principal medio de comunicación entre los distintos participantes en la asignatura.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen	Examen mixto
Preguntas test	10
Preguntas desarrollo	1
Duración del examen	120 (minutos)
Material permitido en el examen	

Calculadora científica.

Criterios de evaluación

La prueba presencial se trata de un cuestionario de 10 preguntas teórico-prácticas que versarán sobre los contenidos de la asignatura y un ejercicio de desarrollo. Cada cuestión tendrá un máximo de cuatro respuestas posibles, siendo sólo correcta una. Cada cuestión tendrá un valor de un punto en caso de contestar de forma correcta, y restará 0.2 puntos en caso de contestarse de forma errónea.

% del examen sobre la nota final 60

Nota del examen para aprobar sin PEC

Nota máxima que aporta el examen a la calificación final sin PEC 4

Nota mínima en el examen para sumar la PEC 4

Comentarios y observaciones

En caso de haber aprobado el examen pero no haber aprobado las prácticas, la nota del examen se guardará para la convocatoria extraordinaria de septiembre en el curso presente.

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad Si

Descripción

La prueba presencial se trata de un cuestionario de 10 preguntas teórico-prácticas que versarán sobre los contenidos de la asignatura y un ejercicio de desarrollo consistente en comentar alguna de las técnicas estudiadas en el curso. Cada cuestión tendrá un máximo de cuatro respuestas posibles, siendo sólo correcta una. Cada cuestión tendrá un valor de un punto en caso de contestar de forma correcta, y restará 0.2 puntos en caso de contestarse de forma errónea. El ejercicio se evaluará hasta un máximo de 5 puntos. Durante la realización de la prueba se podrá utilizar calculadora científica. La prueba presencial se realizará en el Centro Asociado que corresponda a cada estudiante, en las fechas y horarios establecidos por la UNED.

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 no presencial

Descripción

PEC: Actividad 1. 28% de la nota final.

PEC: Actividad 2. 28% de la nota final.

PEC: Actividad 3. 28% de la nota final.

PEC: Actividad 4. 16% de la nota final.

Criterios de evaluación

Los criterios de evaluación se comentan en las descripciones de las actividades que están disponibles en el curso virtual.

Ponderación de la PEC en la nota final 40%

Fecha aproximada de entrega

Comentarios y observaciones

En caso de haber aprobado las PECs pero no haber aprobado el examen, la nota de las PECs se guardará para la convocatoria extraordinaria de septiembre en el curso presente.

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?

La calificación final de la asignatura se calculará de la siguiente forma:

Si en el examen presencial, los trabajos no se obtiene al menos el 40% de la puntuación, entonces el/la estudiante estará suspenso.

En otro caso, se calculará la nota final sumando las diferentes pruebas de evaluación ponderadas con los porcentajes descritos arriba.

Aprobarán la asignatura los estudiantes que consigan al menos 5 puntos en la nota final.

BIBLIOGRAFÍA BÁSICA

ISBN(13):9781119545637

Título:PYTHON MACHINE LEARNING (Abril 2019)

Autor/es:Wei-Meng Lee ;

Editorial:: WILEY

ISBN(13):9781491912058

Título:PYTHON DATA SCIENCE HANDBOOK

Autor/es:Jake Vanderplas ;

Editorial:O'Reilly Media

ISBN(13):9781492032649

Título: HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN, KERAS AND TENSORFLOW
(Second)

Autor/es:Aurélien Géron ;

Editorial:O'Reilly Media

Los libros incluidos en la bibliografía básica se pueden descargar gratuitamente o son accesibles desde la colección de acceso restringido de la Biblioteca de la UNED O'Reilly for Higher Education (New Safari).

BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13):9780387848570

Título:THE ELEMENTS OF STATISTICAL LEARNING (Second)

Autor/es:Tibshirani, Robert J. ; Hastie, Trevor ; Friedman, Jerome ;

Editorial:Springer

ISBN(13):9781788994170

Título: PYTHON MACHINE LEARNING BLUEPRINTS (Second)

Autor/es:Alexander Combs ; Michael Roman ;

Editorial:Packt Publishing

ISBN(13):9781789347999

Título:MACHINE LEARNING ALGORITHMS (Second)

Autor/es:Giuseppe Bonaccorso ;

Editorial:Packt Publishing

ISBN(13):9781789536089

Título:HANDS-ON NEURAL NETWORKS WITH KERAS

Autor/es:Niloy Purkait ;

Editorial:Packt Publishing

Los libros incluidos en la bibliografía complementaria son accesibles desde la colección de acceso restringido de la Biblioteca de la UNED O'Reilly for Higher Education (New Safari).

RECURSOS DE APOYO Y WEBGRAFÍA

Los/as estudiantes dispondrán de los siguientes recursos de apoyo al estudio:

- **Guía de la asignatura.** Incluye el plan de trabajo y orientaciones para su desarrollo. Esta guía será accesible desde el curso virtual.
 - **Curso virtual.** A través de esta plataforma los/as estudiantes tienen la posibilidad de consultar información de la asignatura, realizar consultas al Equipo Docente a través de los foros correspondientes, consultar e intercambiar información con el resto de los compañeros/as.
 - **Biblioteca.** El estudiante tendrá acceso tanto a las bibliotecas de los Centros Asociados como a la biblioteca de la Sede Central, en ellas podrá encontrar un entorno adecuado para el estudio, así como de distinta bibliografía que podrá serle de utilidad durante el proceso de aprendizaje. Además, desde la biblioteca digital de la UNED, el estudiante tendrá acceso a O'Reilly for Higher Education (New Safari), una biblioteca digital con más de 30.000 libros técnicos en constante actualización.
-

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.