



FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Machine Learning II
Código	DOI-MBD-524
Título	Máster en Big Data. Tecnología y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics
Cuatrimestre	Semestral
Créditos	6,0 ECTS
Carácter	Obligatoria
Departamento / Área	Departamento de Organización Industrial
Responsable	Eugenio F. Sánchez Úbeda
Horario de tutorías	Fijar cita previa por email
Datos del profesorado	
Profesor	
Nombre	Alberto Gascón González
Departamento / Área	Departamento de Organización Industrial
Correo electrónico	alberto.gascon@icai.comillas.edu
Profesor	
Nombre	Álvaro Jesús López López
Departamento / Área	Instituto de Investigación Tecnológica (IIT)
Despacho	Francisco de Ricci, 3
Correo electrónico	Alvaro.Lopez@iit.comillas.edu
Teléfono	4503
Profesor	
Nombre	Eugenio Francisco Sánchez Úbeda
Departamento / Área	Departamento de Organización Industrial
Despacho	Santa Cruz de Marcenado 26
Correo electrónico	Eugenio.Sanchez@iit.comillas.edu
Teléfono	2706
Profesor	
Nombre	Jorge Ayuso Rejas
Departamento / Área	Departamento de Ciencias de la Computación e Inteligencia Artificial (DCIA)
Correo electrónico	jayuso@icai.comillas.edu
Profesor	



Nombre	Roberto Kramer Neves Pinto
Departamento / Área	Escuela Técnica Superior de Ingeniería (ICAI)
Correo electrónico	rkramer@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura

Aportación al perfil profesional de la titulación

Este curso se centra en un conjunto de herramientas de aprendizaje automático para modelar y comprender conjuntos de datos complejos. En particular, el curso describe las herramientas seleccionadas agrupadas según tres enfoques principales: métodos de ensamblado, aprendizaje profundo y aprendizaje por refuerzo. Estos enfoques principales se complementan con otros temas como algoritmos genéticos y sistemas de recomendación. Para lograr este objetivo, el curso se ha planteado como un equilibrio entre la teoría y la práctica.

Prerrequisitos

Se requieren conocimientos básicos de Cálculo y Álgebra (comprender y manipular ecuaciones, comprensión total de funciones y funciones inversas, comprender límites, derivadas e integrales, conocer reglas para producto y suma, etc.). También se requieren conocimientos básicos de Estadística (estadística descriptiva, modelos de distribución de probabilidad discretos y continuos, muestreo y conceptos básicos de inferencia estadística), así como conocimientos sobre métodos estándar de Machine Learning (regresión, clasificación y aprendizaje no supervisado).

Se requieren conocimientos básicos de Programación en R y Python para las sesiones de práctica.

Competencias - Objetivos

Competencias

AAAConocimientos o contenidos

- C01 Entender los fundamentos de la analítica de datos y su aplicación en diversas áreas de la inteligencia artificial, destacando la integración en soluciones complejas y multidisciplinares para el análisis avanzado de datos masivos atendiendo a la diversidad de problemas específicos de cada área.
- C02 Comprender las técnicas de procesados de datos, las arquitecturas y herramientas más habituales y apropiadas para condiciones y requisitos de casos específicos
- C03 Comprender las técnicas de machine Learning e Inteligencia artificial, desde las más sencillas a las más sofisticadas y ser capaz de seleccionar la más adecuada e implementarla de una manera adecuada a las condiciones y requisitos del caso específico.

AAACompetencias

CP1	Integrar las arquitecturas, técnicas de inteligencia artificial, análisis avanzado de datos y de visualización y de cumplimiento legal para ofrecer la solución global óptima.
CP3	Implementar técnicas de análisis estadístico y lenguajes de programación para Machine Learning en el contexto de Big Data, ajustando la metodología a las especificidades de cada conjunto de datos para optimizar los resultados obtenidos.



CP4 Implementar las técnicas de procesado de datos y usar las herramientas más habituales y apropiadas a las condiciones y requisitos de casos específicos.

CP5 Aplicar los principios éticos relativos a la recogida, almacenamiento, y análisis de datos teniendo en cuenta las posibles discriminaciones directas o indirectas derivadas de la toma de decisiones.

CP7 Aplicar conocimientos avanzados en Big Data y analítica de datos para desarrollar soluciones innovadoras en proyectos y en investigación, aportando y evaluando soluciones óptimas para el procesamiento y análisis de datos a gran escala.

AAAHabilidades o destrezas

HA1 Comunicar de manera oral y escrita con rigor técnico, claridad expositiva y coherencia argumentativa a todo tipo de interlocutores, técnicos y no técnicos.

HA2 Trabajar en equipos de carácter pluridisciplinar y/o internacional y organizar y liderar adecuadamente las dinámicas de grupo.

HA3 Desarrollar las habilidades interpersonales que requieren los entornos profesionales actuales (empatía, tolerancia, respeto, capacidad para aunar intereses contrapuestos).

HA4 Gestionar, organizar y planificar adecuadamente el trabajo y el tiempo, cumpliendo objetivos y estándares de calidad.

HA5 Mantener una formación y aprendizaje continuo y adaptación a los cambios tecnológicos y científicos.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Bloque 1: Ensamblado de modelos

- Aprendizaje mediante combinación de modelos. Sesgo y varianza de un modelo.
- Ensamblado de modelos mediante agregación. Bootstrap para generar aleatoriedad de modelos. Reducción de la varianza mediante Bagging. Random Forests como mejora de Bagging. Ajuste de hiperparámetros.
- Ensamblado de modelos mediante Boosting. Ajuste incremental de modelos aditivos. Descenso del gradiente y Gradient Boosting. Ajuste de hiperparámetros.

Bloque 2: Algoritmos genéticos

- Aprendizaje y Técnicas evolutivas en IA. Fundamentos de Algoritmos genéticos.

Bloque 3: Aprendizaje profundo

- Redes neuronales convolucionales (CNN): origen, capas y arquitecturas principales.
- Entrenamiento de redes neuronales profundas: funciones de pérdida, descenso del gradiente, propagación inversa (backpropagation), tasa de aprendizaje y regularización. Aumento de datos (data augmentation).
- Aprendizaje profundo no supervisado. Taxonomía de modelos generativos. Fundamentos de autocodificadores (autoencoders) y redes generativas antagónicas (GAN).

Bloque 4: Aprendizaje por refuerzo

- Procesos de decisión Markovianos. Ecuación de Bellman.
- Métodos tabulares de aprendizaje por refuerzo. Conceptos de aprendizaje on-policy y off-policy.
- Métodos no tabulares. Enfoque DQN y derivados.

Bloque 5: Sistemas de recomendación

- Fundamentos de los sistemas de recomendación. Reglas de asociación y filtrado colaborativo.



METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

La metodología utilizada se centrará en facilitar el aprendizaje tanto de los conceptos teóricos como la puesta en práctica de los mismos, requiriendo la participación activa de los estudiantes. Además, las actividades realizadas en clase de forma presencial se deben complementar con el trabajo individual del estudiante fuera del horario de clase.

Metodología Presencial: Actividades

Clases magistrales expositivas y participativas: El profesor introducirá los conceptos fundamentales de cada tema, junto con algunas recomendaciones prácticas, y pasará por ejemplos ilustrativos para apoyar la explicación. Se incentivará la participación activa planteando preguntas abiertas para fomentar la discusión

Sesiones prácticas con uso de software: Los estudiantes trabajarán bajo la supervisión del profesor, aplicando los conceptos y técnicas descritos en clase a problemas reales.

Metodología No presencial: Actividades

Estudio personal: Los estudiantes revisarán el material disponible para entender e interiorizar los conceptos teóricos de la asignatura y aprender a ponerlos en práctica.

Elaboración de trabajos: Los estudiantes realizarán ejercicios prácticos con el ordenador para fijar los conceptos teóricos de la asignatura y desarrollar las habilidades necesarias para poder resolver diferentes tipos de problemas utilizando las técnicas consideradas.

RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES	
Teoría	35
Práctica	25
HORAS NO PRESENCIALES	
Estudio	40
Realización de trabajos colaborativos	60
Trabajo autónomo sobre contenidos prácticos	20

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	Peso
Exámenes: <ul style="list-style-type: none">Exámenes de carácter teórico-práctico.	<ul style="list-style-type: none">Comprensión de conceptos teóricos.Aplicación de los conceptos a la resolución de problemas.Ánalysis e interpretación de los resultados.Comparación de resultados utilizando diferentes técnicas.Uso de software estadístico.	50



Actividades de evaluación	Criterios de evaluación	Peso
Evaluación continua del rendimiento: <ul style="list-style-type: none">Trabajos de carácter práctico individual o en grupo.Ejercicios o problemas resueltos de manera individual o en grupo.Pruebas cortas de evaluación continua.Participación en clase.Asistencia y actitud en clase.	<ul style="list-style-type: none">Comprensión de conceptos teóricos.Aplicación de los conceptos a la resolución de problemas.Análisis e interpretación de los resultados.Comparación de resultados utilizando diferentes técnicas.Uso de software estadístico.	50

Calificaciones

Se deben cumplir las siguientes condiciones para aprobar el curso:

- Una calificación global mínima de al menos 5 sobre 10.
- Una nota mínima en el examen final de 4 sobre 10.

La calificación global se obtiene de la siguiente manera:

- El examen final representa el 35% de la calificación final si la calificación en este examen es al menos 4. En otro caso, el examen final representa el 100% de la calificación general.
- El examen parcial representa el 15%.
- Los trabajos prácticos suponen el 50% de la nota final.

PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
Introducción	Semana 1	
Ensamblado modelos (teoría)	Semana 1	
Ensamblado modelos (teoría)	Semana 2	
Ensamblado modelos (práctica)	Semana 2	
Ensamblado modelos (teoría)	Semana 3	
Ensamblado modelos (teoría)	Semana 3	
Ensamblado modelos (práctica)	Semana 4	



Algoritmos genéticos (teoría)	Semana 4	
Algoritmos genéticos (teoría)	Semana 5	
Algoritmos genéticos (práctica)	Semana 5	
Backfitting	Semana 6	
Examen Parcial	Semana 6	
Sistemas de recomendación (teoría)	Semana 7	
Sistemas de recomendación (práctica)	Semana 7	
Aprendizaje profundo (teoría)	Semana 8	
Aprendizaje profundo (teoría)	Semana 8	
Aprendizaje profundo (teoría)	Semana 9	
Aprendizaje profundo (práctica)	Semana 9	
Aprendizaje profundo (teoría)	Semana 10	
Aprendizaje profundo (práctica)	Semana 10	
Aprendizaje profundo (práctica)	Semana 11	
Aprendizaje profundo (práctica)	Semana 11	
Aprendizaje profundo (práctica)	Semana 12	
Aprendizaje por refuerzo (teoría)	Semana 12	
Aprendizaje por refuerzo (teoría)	Semana 13	
Aprendizaje por refuerzo (práctica)	Semana 13	
Aprendizaje por refuerzo (teoría)	Semana 14	
Aprendizaje por refuerzo (práctica)	Semana 14	



BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Goodfellow, I., Bengio, Y. and Aaron Courville, A. (2016). Deep Learning. Sixth Edition. Springer.
- Bishop, C. M. & Bishop, H. (2023). *Deep Learning: Foundations and Concepts*. Springer International Publishing AG. ISBN 978-3-031-45467-7.
- James, G., Witten, D., Hastie, T. and Tibshirani, R. (2023). Introduction to Statistical Learning with Applications in Python. Sixth Edition. Springer (<https://www.statlearning.com>)
- Hastie, T., Tibshirani, R., Friedman, J. (2017). The Elements of Statistical Learning: Data Mining, Inference and Prediction. Second Edition, Springer.

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedeelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>