Aprendizaje automático II

MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS / MASTER IN DATA SCIENCE

UNIVERSIDAD INTERNACIONAL MENÉNDEZ PELAYO

Másteres Universitarios

Este documento puede utilizarse como documentación de referencia de esta asignatura para la solicitud de reconocimiento de créditos en otros estudios. Para su plena validez debe estar sellado por la Secretaría de Estudiantes UIMP.

DATOS GENERALES

Breve descripción

Los objetivos de esta asignatura son:

- Conocer los fundamentos del aprendizaje estadístico.
- Conocer los métodos de aprendizaje automático basados en funciones núcleo (metodos kernel).
- Saber aplicar los métodos kernel en problemas prácticos de clasificación, regresión y análisis de datos.
- Conocer el uso de variables latentes en modelo gráficos y probabilísticos.
- Saber aplicar modelos y redes probabilísticas en problemas de análisis de datos y aprendizaje automático.

Título asignatura

Aprendizaje automático II

Código asignatura

102269

Curso académico

2025-26

Planes donde se imparte

MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS / MASTER IN DATA SCIENCE

Créditos ECTS

4

Carácter de la asignatura

OPTATIVA

Duración

Cuatrimestral

Idioma

Castellano e Inglés

CONTENIDOS

Contenidos

1. Aprendizaje estadístico.
2. Márgenes y vectores soporte. Máquinas de vector soporte (SVM).
3. Métodos basados en núcleos.
4. Variables latentes y método EM.
5. Modelos de Markov ocultos (HMM).
6. Aprendizaje Bayesiano. Redes probabilísticas. Causalidad.
7. Selección de modelos. MCMC.

RESULTADOS DE APRENDIZAJE Y DE FORMACIÓN

Generales

CG1 - Capacidad para integrarse eficazmente en un grupo de trabajo y trabajar en equipo, compartir la información disponible e integrar su actividad en la actividad del grupo colaborando de forma activa en la consecución de objetivos comunes

CG2 - Capacidad de estudio, síntesis y autonomía suficientes para, una vez finalizado este programa formativo, iniciar una Tesis Doctoral

CG4 – Saber preparar y conducir presentaciones, ante públicos especializado, sobre una investigación o proyecto científico

CG7 - Conocer las herramientas metodológicas necesarias para desarrollar proyectos avanzados

CG8 - Capacidad de actualización de los conocimientos expuestos en el ámbito de la comunidad científica

Transversales

CT1 - Capacidad para buscar, obtener, seleccionar, tratar, analizar y comunicar información utilizando diferentes fuentes

CT2 - Capacidad para proyectar los conocimientos, habilidades y destrezas adquiridos para promover una sociedad basada en los valores de la libertad, la justicia, la igualdad y el pluralismo

CT3- Dominio de la gestión del tiempo

CT4- Capacidad para afrontar tareas y situaciones críticas

CT5- Capacidad de trabajo autónomo y toma de decisiones

CT6- Capacidades asociadas al trabajo en equipo: cooperación, liderazgo, saber escuchar

Específicas

DSDA01 - Utilizar el análisis predictivo para analizar grandes volúmenes de datos y descubrir nuevas relaciones

DSDA02 - Utilizar técnicas estadísticas apropiadas sobre los datos disponibles para lograr una visión adecuada de los mismos

DSENG02 - Desarrollar y aplicar soluciones computacionales para problemas en un cierto

dominio de aplicación, usando una amplia gama de plataformas de análisis de datos

DSRM06 - Aplicar el ingenio propio para resolver problemas complejos y desarrollar ideas innovadoras

PLAN DE APRENDIZAJE

Actividades formativas

AF1 - Participación y asistencia a lecciones magistrales y seminarios (15 horas)

AF2 - Realización de prácticas de computación y análisis de datos (15 horas)

AF6 - Tutorías (presenciales o por medio de recursos telemáticos) (8 horas)

AF7 - Elaboración de informes de laboratorio y trabajos (20 horas)

AF8 - Estudio individual de contenidos de la asignatura (20 horas)

AF9 - Trabajo en grupo (20 horas)

A10 - Pruebas de evaluación (2 horas)

Metodologías docentes

Se comenzará por una exposición de los conceptos y métodos básios, incluyendo ejemplos ilustrativos sencillos de distintas disciplinas, que serñan analizados y discutidos en común.
Los estudiantes, organizados en grupos, realizarán un análisis práctico detallado de distintos casos de estudio relaes (por ejemplo, problemas de Kaggle o Kaggle in Class) adaptando y aplicando las técnicas adecuadas en cada caso.
La parte práctica se realizará utilizando el software estadístico R de uso libre y software específico especializado (e.g.: deep learning).

Resultados de aprendizaje

- Entender los fundamentos del aprendizaje estadístico.
- Entender el concepto de variable latente y conocer los métodos para su tratamiento.
- Saber analizar datos discretos utilizando redes probabilísticas.

SISTEMA DE EVALUACIÓN

Descripción del sistema de evaluación

SE1 - Examen (escrito, oral y/o práctico en el aula de computación) (40%)

SE2 - Valoración de informes y trabajos escritos (60%)

PROFESORADO

Profesor responsable

Santamaría Caballero, Luis Ignacio

Doctor Ingeniero de Telecomunicación.

Catedrático de Universidad (Teoría de la Señal y Comunicaciones).

Universidad de Cantabria.

Profesorado

Herrera García, Sixto

Doctor por la Universidad de Cantabria.

Programa: Matemáticas y Computación.

Profesor Titular de Universidad (Matemática Aplicada).

UC. Dpto. de Matemática Aplicada y Ciencias de la Computación (Caminos).

Van Vaerenbergh , Steven Johan Maria

Doctor por la Universidad de Cantabria.

Profesor Contratado Doctor.

Universidad de Cantabria.

Cuevas Fernández, Diego

Doctorado en Tecnologías de la Información y Comunicaciones en Redes Móviles.

Contratado de Investigación en Formación (Predoctoral FPU).

Dpto. Ingeniería de Comunicaciones, Universidad de Cantabria.

Legasa Ríos, Mikel Néstor

Máster en Matemáticas y Computación.

Contratado Investigación, Dpto. Matemática Aplicada y Ciencias de la Computación. (UC)

HORARIO

Horario

02/02/2026

15:30 - 17:30

SVMs. Clasificación

Luis Ignacio Santamaría Caballero

Doctor Ingeniero de Telecomunicación.
Catedrático de Universidad (Teoría de la Señal y Comunicaciones).
Universidad de Cantabria.

04/02/2026

15:30 - 16:30

Práctica clasificación

Steven Johan Maria Van Vaerenbergh

Doctor por la Universidad de Cantabria.
Profesor Contratado Doctor.
Universidad de Cantabria.

16:30 - 17:30

Práctica clasificación

Diego Cuevas Fernández

Doctorado en Tecnologías de la Información y Comunicaciones en Redes Móviles.
Contratado de Investigación en Formación (Predoctoral FPU).

Dpto. Ingeniería de Comunicaciones, Universidad de Cantabria.

09/02/2026

15:30 - 17:30

Métodos kernel para regresión. SVR. Kernel Ridge Regression. Gaussian processes.

Luis Ignacio Santamaría Caballero

Doctor Ingeniero de Telecomunicación.
Catedrático de Universidad (Teoría de la Señal y Comunicaciones).
Universidad de Cantabria.

11/02/2026

15:30 - 16:30

Práctica regresión

Steven Johan Maria Van Vaerenbergh

Doctor por la Universidad de Cantabria.
Profesor Contratado Doctor.
Universidad de Cantabria.

16:30 - 17:30

Práctica regresión

Diego Cuevas Fernández

Doctorado en Tecnologías de la Información y Comunicaciones en Redes Móviles.
Contratado de Investigación en Formación (Predoctoral FPU).

Dpto. Ingeniería de Comunicaciones, Universidad de Cantabria.

16/02/2026

15:30 - 17:30

Aproximación de kernels. Métodos kernel escalables. Aprendizaje online. Métodos kernel adaptativos.

Luis Ignacio Santamaría Caballero

Doctor Ingeniero de Telecomunicación.
Catedrático de Universidad (Teoría de la Señal y Comunicaciones).
Universidad de Cantabria.

18/02/2026

15:30 - 16:30

Práctica aprendizaje online/large-scale

Steven Johan Maria Van Vaerenbergh

Doctor por la Universidad de Cantabria.
Profesor Contratado Doctor.
Universidad de Cantabria.

16:30 - 17:30

Práctica aprendizaje online/large-scale

Diego Cuevas Fernández

Doctorado en Tecnologías de la Información y Comunicaciones en Redes Móviles.

Contratado de Investigación en Formación (Predoctoral FPU).
Dpto. Ingeniería de Comunicaciones, Universidad de Cantabria.

23/02/2026

15:30 - 17:30

Métodos kernel no supervisados. Kernel PCA. Spectral Clustering. Kernel k-medias.

Luis Ignacio Santamaría Caballero

Doctor Ingeniero de Telecomunicación.
Catedrático de Universidad (Teoría de la Señal y Comunicaciones).
Universidad de Cantabria.

25/02/2026

15:30 - 16:30

Práctica clustering

Steven Johan Maria Van Vaerenbergh

Doctor por la Universidad de Cantabria.
Profesor Contratado Doctor.
Universidad de Cantabria.

16:30 - 17:30

Práctica clustering

Diego Cuevas Fernández

Doctorado en Tecnologías de la Información y Comunicaciones en Redes Móviles.

Contratado de Investigación en Formación (Predoctoral FPU).
Dpto. Ingeniería de Comunicaciones, Universidad de Cantabria.

03/03/2026

15:30 - 17:30

Redes probabilísticas discretas

Sixto Herrera García

Doctor por la Universidad de Cantabria. Programa: Matemáticas y Computación.
Profesor Titular de Universidad (Matemática Aplicada).
Universidad de Cantabria. Dpto. de Matemática Aplicada y Ciencias de la Computación (Caminos).

05/03/2026

15:30 - 17:30

Práctica de redes bayesianas: creación e inferencia

Maialen Iturbide Martínez de Albéniz

Investigadora Doctora en Agrobiología Ambiental
Universidad de Cantabria (UC)

10/03/2026

15:30 - 17:30

Aprendizaje Estructural

Sixto Herrera García

Doctor por la Universidad de Cantabria. Programa: Matemáticas y Computación.
Profesor Titular de Universidad (Matemática Aplicada).
Universidad de Cantabria. Dpto. de Matemática Aplicada y Ciencias de la Computación (Caminos).

12/03/2026

15:30 - 17:30

Aprendizaje Estructural

Maialen Iturbide Martínez de Albéniz

Investigadora Doctora en Agrobiología Ambiental
Universidad de Cantabria (UC)

17/03/2026

15:30 - 17:30

Aprendizaje paramétricos/Redes Gaussianas Mixtas

Sixto Herrera García

Doctor por la Universidad de Cantabria. Programa: Matemáticas y Computación.
Profesor Titular de Universidad (Matemática Aplicada).
Universidad de Cantabria. Dpto. de Matemática Aplicada y Ciencias de la Computación (Caminos).

19/03/2026

15:30 - 17:30

Clasificadores Bayesianos. Naive Bayes.

Maialen Iturbide Martínez de Albéniz

Investigadora Doctora en Agrobiología Ambiental
Universidad de Cantabria (UC)

24/03/2026

15:30 - 17:30

Prácticas: Aprendizaje

Maialen Iturbide Martínez de Albéniz

Investigadora Doctora en Agrobiología Ambiental
Universidad de Cantabria (UC)

26/03/2026

15:30 - 16:30

Examen

Luis Ignacio Santamaría Caballero

Doctor Ingeniero de Telecomunicación.
Catedrático de Universidad (Teoría de la Señal y Comunicaciones).
Universidad de Cantabria.

16:30 - 17:30

Examen

Sixto Herrera García

Doctor por la Universidad de Cantabria. Programa: Matemáticas y Computación.
Profesor Titular de Universidad (Matemática Aplicada).
Universidad de Cantabria. Dpto. de Matemática Aplicada y Ciencias de la Computación

(Caminos).

BIBLIOGRAFÍA Y ENLACES RELACIONADOS

Bibliografía

Bibliografía básica

S. Y. Kung, *Kernel Methods and Machine Learning*, Cambridge University Press, 2014

B. Schölkopf, A. J. Smola, *Learning with Kernels*, The MIT Press, 2002

C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006

K. P. Murphy, *Machine Learning: A Probabilistic Perspectiv*e, The MIT Press 2012

R. Nagarajan, M. Scutari, S. Lèbre, *Bayesian Networks in R with Applications in Systems Biology*, Springer, 2013

M. Scutari, J.B. Denis, *Bayesian Networks with examples in R*, CRC Press, 2014

E. Castillo, J.M. Gutiérrez, A.S. Hadi, *Expert Systems and Probabilistic Network Mode*ls, Springer-Verlag, 1997

Bibliografía complementaria

J. Shawe-Taylor, N. Cristianini, *Kernel Methods for Pattern Analysi*s, Cambridge University Press, 2004

C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press 2006