

# Aprendizaje automático (AA)

Alberto Díaz  
Curso 2022-2023



# Tutorías y comunicación

- Tutorías (despacho 445)
  - X: 10-13
  - enviar correo antes
- Campus Virtual
  - APRENDIZAJE AUTOMÁTICO y BIG DATA
- Comunicación de la asignatura
  - A través del campus virtual: Correo/mensaje

# Programa

1. Introducción al aprendizaje automático.
2. Métodos de regresión.
3. Redes neuronales.
4. Evaluación de sistemas de aprendizaje automático.
5. Árboles de decisión
6. Aprendizaje no supervisado.
7. Detección de anomalías.
8. Sistemas de recomendación.

# Evaluación

- Ejercicios en cada tema
- 3 prácticas, aproximadamente
  - Es obligatorio entregar las prácticas
- La evaluación, tanto en la convocatoria ordinaria como en la extraordinaria, se basa en un trabajo, aplicando alguna(s) de las técnica(s) desarrolladas en las prácticas a un conjunto de datos de vuestra elección
  - Defensa del proyecto: 20% de la nota
  - Trabajo escrito sobre el proyecto: 80% de la nota

# Proyecto

- El objetivo del proyecto es trabajar sobre un problema de clasificación con un conjunto de datos que hay que seleccionar en la plataforma Kaggle (<https://www.kaggle.com/datasets>)
  - Hay que enviar una propuesta al profesor de la asignatura (enviando un correo con un enlace al conjunto de datos propuesto) para que dé su visto bueno.
  - Has de utilizar el código de las prácticas que has hecho en la asignatura para desarrollar un sistema de aprendizaje automático sobre el conjunto de datos en cuestión.
  - Se debe dar una estimación de la efectividad del sistema, comparando distintas técnicas de aprendizaje automático estudiadas en la asignatura.
  - Se recomienda aplicar regresión logística, redes neuronales y árboles de decisión.
  - Adicionalmente, se pueden también aplicar implementaciones y técnicas no desarrolladas en la asignatura.
  - No puede haber dos proyectos que utilicen el mismo conjunto de datos.

# Dinámica

- Teoría los viernes
- Laboratorio los miércoles, a partir del 9 de febrero
  - 48 alumnos matriculados
    - laboratorio 10: 25 puestos
    - laboratorio 11: 25 puestos
- Tanto las prácticas como los proyectos se realizarán en grupos de 2 o 3 alumnos
  - Cada grupo me debe comunicar los integrantes que lo forman (campus virtual)

# Bibliografía

- Coursera.
  - <https://www.coursera.org/>
    - [Programa especializado: Aprendizaje automático](#)
- Python
  - Multitud de recursos online.
  - <https://www.kaggle.com/learn>
  - <https://scipy-lectures.org/>
  - <https://cursosinformatica.ucm.es/>
    - 2 Créditos ECTS optativos de grado
    - Introducción a la programación en Python
    - Análisis de datos con Python
    - Visualización de datos con Python

# Bibliografía

- Aurélien Géron; Hands-On Machine Learning with Scikit-Learn and TensorFlow, 2nd edition; O'Reilly Media, 2019
- Scipy Lecture Notes: <https://scipy-lectures.org>
- Wes McKinney; Python for Data Analysis, 2nd Edition; O'Reilly Media, 2017
- Sebastian Raschka, Vahid Mirjalili; Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow, 2nd Edition; Packt Publishing, 2017
- Ian H. Witten, Eibe Frank, Mark A. Hall and Christopher J. Pal; Data Mining: Practical Machine Learning Tools and Techniques, Fourth Edition; Morgan Kaufmann, 2016
- Tom M. Mitchell; Machine Learning; McGraw-Hill, 1997



# Software

- Anaconda
  - <https://www.anaconda.com/products/individual>
  - Labs:
    - Jupyter Notebook - Anaconda 3
    - Google Collab