



## Módulo R - Actividad 5. Videoconferencia final

MACHINE LEARNING E INTELIGENCIA  
ARTIFICIAL ONLINE  
*On-line* -

Curso académico 2021

**PROFESOR**

Javier Abascal Carrasco



Esta publicación está bajo licencia Creative Commons Reconocimiento, No comercial, Compartir igual, (by-nc-sa). Usted puede usar, copiar y difundir este documento o parte del mismo siempre y cuando se mencione su origen, no se use de forma comercial y no se modifique su licencia. Más información: <http://creativecommons.org/licenses/by-nc-sa/3.0/>

## Índice

### 1. Machine Learning con R. Clasificación

# 1. Machine Learning con R. Clasificación

Durante esta sesión vamos a realizar el ejercicio de clasificación del script de R “Machine Learning con R (segunda parte).R”

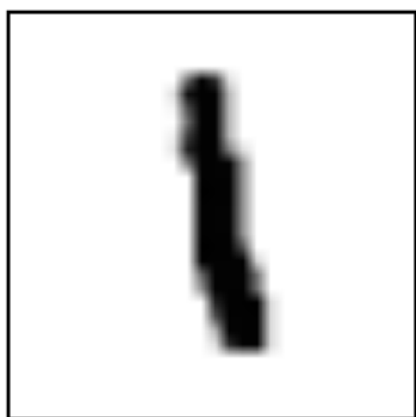
En este ejercicio trabajaremos con el famoso dataset de MNIST

Source: <http://yann.lecun.com/exdb/mnist/>

*"The MNIST database of handwritten digits, available from this page, has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The digits have been size-normalized and centered in a fixed-size image."*

Cada imagen está formada por una matriz de 28x28 pixels y cada pixel es una escala de grises con el valor desde 0 (blanco) a 255 (negro)

Para trabajar con imágenes de forma sencilla, se suele realizar un “flatten” aunque se pierda información local del mismo. Es decir, se trabaja con una fila de  $28 \times 28 = 784$  columnas



21

[illegible]

Durante el Webinar se explicarán los conceptos de forma intuitiva, pero no habrá tiempo de profundizar en todos los métodos y algoritmos utilizados. Por ello, a continuación se facilitan una serie de links para leer y entender bien el ejercicio que realizaremos de clasificación (leer únicamente la sección que se dice). Si eres una persona que ya estás familiarizado con los términos, no tienes que hacer nada, si no estás acostumbrado a los términos de machine learning, aconsejo leer todos los enlaces de abajo antes del webinar - con el fin de sacar el máximo provecho al mismo

- Introducción a árboles de decisión como métodos de Machine Learning - [aquí](#)
- Entendimiento general de un árbol de regresión y un árbol de clasificación - obviar la parte matemática si es necesario
- El algoritmo de [Random Forest](#) para clasificación
- El algoritmo de [GBM](#) (Gradient Boosting Machine) para clasificación

Adicionalmente, trabajaremos un ejercicio de clustering

- Qué es el [clustering](#)
- [Kmeans](#) - existen muchos vídeos explicativos
- Reducción de [dimensionalidad con PCA](#) - Principal Component Analysis
- Reducción de [dimensionalidad con TSNE](#) - (t-Distributed Stochastic Neighbor Embedding)