

# *Deep learning*

Fabián Villena

Enero 2024

Se le pide detectar la malignidad de una muestra histológica de un paciente en función de parámetros morfológicos de los núcleos de las células <sup>1</sup>.

En particular, trabajaremos con un conjunto de datos que contiene múltiples atributos morfológicos de núcleos celulares de células de una muestra histológica de un tejido tumoral de un paciente. Debe predecir la columna *diagnosis*, la cual puede pertenecer a la clase benigno (*B*) o maligno (*M*).

Los resultados esperados son que apliquen una técnica basada en *Deep Learning* para clasificar la malignidad del tumor dadas las múltiples características nucleares de las células. El conjunto de datos se encuentra disponible en la siguiente dirección:

<https://github.com/fvillena/biocompu/blob/2023/data/wisconsin.csv>

## Preguntas

Responda las siguientes preguntas en un *Jupyter Notebook* con código desarrollado en el lenguaje de programación Python.

1. Verifique y preprocese si es necesario las características.
2. Separe el conjunto de datos en un subconjunto de entrenamiento y prueba
3. Diseñe una red neuronal artificial del tipo Totalmente Conectada que resuelva la tarea.
4. Entrene la red neuronal artificial con los datos del subconjunto de entrenamiento.
5. Prediga sobre el subconjunto de prueba utilizando la red neuronal artificial ajustada.
6. Calcule al menos dos métricas de rendimiento para el modelo.

---

<sup>1</sup><https://doi.org/10.1117/12.148698>