

## Trabajo Práctico 02: Redes Neuronales

**Objetivo:** Crear y entrenar una red Perceptrón Multicapa con *BackPropagation* utilizando la biblioteca de Python [TensorFlow](#). Los datos para entrenar la red son provistos por la cátedra y se especifican más abajo.

**Forma de entrega:** El trabajo se deberá entregar por e-mail y deberá incluir:

- El archivo “.py” que corresponde a la red neuronal, escrita en código python.
- El resultado obtenido (sobrevivió o no, 1/0) para cada uno de los 100 pasajeros en el conjunto de prueba (test.csv) indicado en la planilla de Google (hay una columna por grupo):

<https://docs.google.com/spreadsheets/d/1jvqlUGf19I7ayLkrvriZYcFi8iDKR-BHs3tP9Clnmfo/edit?usp=sharing>

- Un informe breve que indique:
  - Arquitectura perceptrón creada
  - Resultados obtenidos:
    - % de casos de éxito (usando el 33% del conjunto de entrenamiento para validar la red)
    - Error total
    - Iteraciones

**Fecha de entrega:** deberá entregarse a más tardar el día 30 de mayo.

### Datos:

Los datos podrán ser descargados aquí:

<https://drive.google.com/file/d/1ngdTupOn2dVqj77kgs-pTgl1F9GsT8co/view?usp=sharing>

Incluye dos archivos de tipo CSV:

- **Train.csv:** será utilizado para entrenar la red y probarla también.
- **Test.csv:** será utilizado para probar la red con casos desconocidos y se compararán los resultados obtenidos por cada grupo.

Básicamente los datos, son una lista en formato CSV de los pasajeros del Titanic, se indican algunos datos de ellos (ver diccionario de datos) y si sobrevivió o no.

La idea es entrenar una red neuronal que permita predecir si un pasajero sobrevivirá o no. Para ello partir el conjunto de entrenamiento (train.csv) en 2. Dejando el 66% de los datos para entrenar la red y el restante 33% para probarla y calcular cuan preciso es este método.

El segundo conjunto de datos: "test.csv", no tiene un el campo "survival", así que este conjunto se utilizará para calcular la supervivencia de cada pasajero y se la indicará en la planilla de Google ya mencionada.

### Diccionario de datos:

| Variable | Definición                                   | Valores posibles                               |
|----------|--|--|
| survival | Si sobrevivió o no                           | 0 = No, 1 = Yes                                |
| pclass   | Clase de ticket, primera clase, etc.         | 1 = 1st, 2 = 2nd, 3 = 3rd                      |
| sex      | Sexo   |  |
| Age      | Edad en años                                 |  |
| sibsp    | # de hermanos / cónyuges a bordo del Titanic |  |
| parch    | # of padres / hijos en el Titanic            |  |
| ticket   | Número de ticket                             |  |
| fare     | tarifa                                       |  |
| cabin    | Número de cabina                             |  |
| embarked | Puerto de embarque                           | C = Cherbourg, Q = Queenstown, S = Southampton |

### Notas sobre las variables:

**pclass:** Un representante del estatus socio-económico

1st = Upper

2nd = Middle

3rd = Lower

**age:** La edad es una fracción si es menor que 1. La edad es estimada si tiene forma de xx.5

**sibsp:** El conjunto de datos define las relaciones familiares de esta manera ...

Sibling = hermano, hermana, hermanastro, hermanastra

Spouse = esposo, esposa (las amantes y los novios fueron ignorados)

**parch:** El conjunto de datos define las relaciones familiares de esta manera ...

Parent = madre, padre

Child = hija, hijo, hijastra, hijastro

Algunos niños viajaron solo con una niñera, por lo tanto parch = 0 para ellos.