

# Minería de Datos para el Análisis de Big Data

Por: Carlos Carreño  
[ccarrenovi@gmail.com](mailto:ccarrenovi@gmail.com)

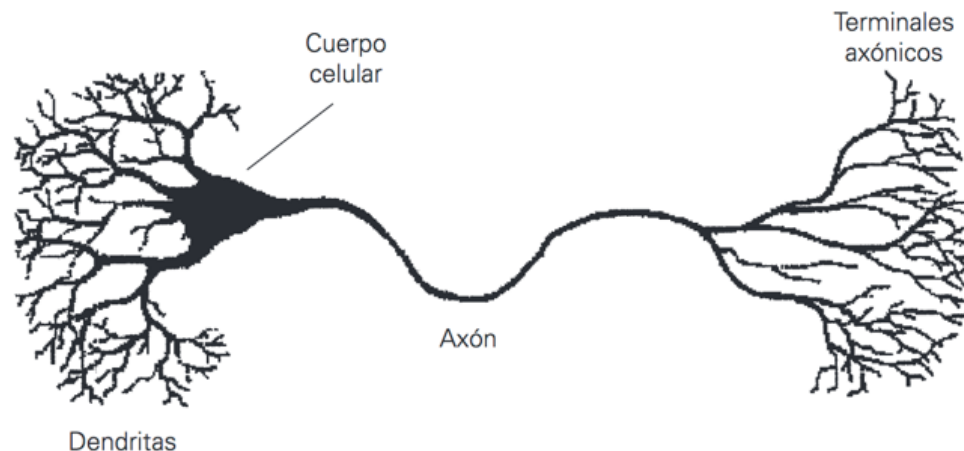
Abril, 2021

# Modulo 10 Redes Neuronales Artificiales

- Introducción
- Neurona Artificial
- Red Neuronal Artificial (RNA)
- Fases de Modelización
- Función de Activación
- Tamaño de las RNA
- Capas Ocultas
- Regresión y las RNA
- Modelo Lineal de las RNA

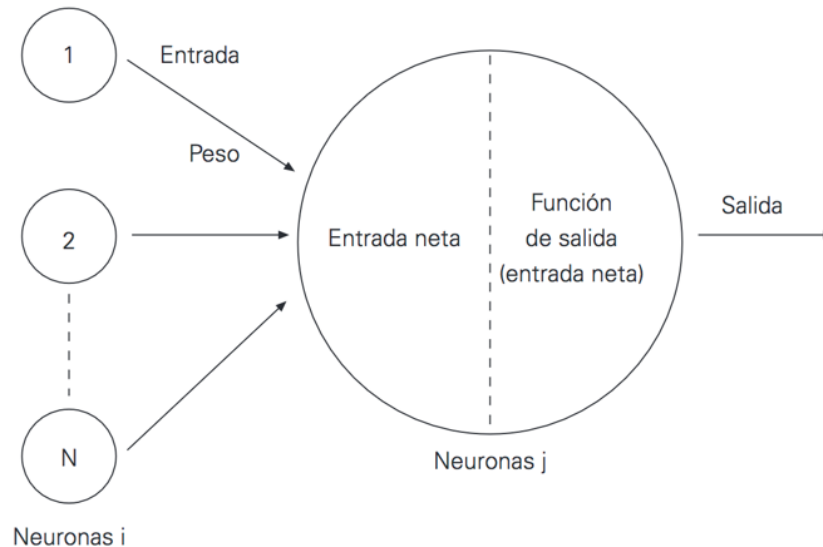
# Introducción a las Redes Neuronales Artificiales

- Las redes neuronales artificiales se inspiran en el comportamiento conocido del cerebro humano (principalmente el referido a las neuronas y sus conexiones), trata de crear modelos artificiales que solucionen problemas difíciles de resolver mediante técnicas algorítmicas convencionales.
- La idea central es extraer las combinaciones lineales de las variables de entrada como características derivadas y, a continuación, modelar el objetivo como una función no lineal de estas características.
- la neurona artificial pretende mimetizar las características más importantes de la neurona biológica.



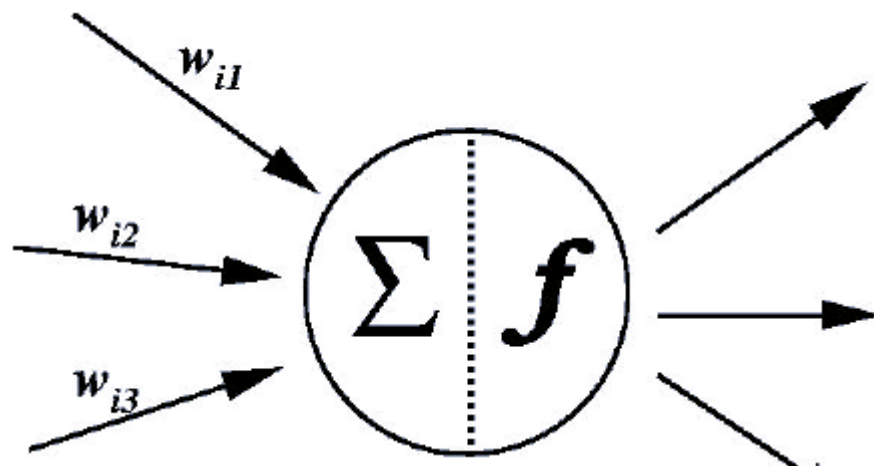
# Neurona Artificial

- En general, recibe las señales de entrada de las neuronas vecinas ponderadas por los pesos de las conexiones. La suma de estas señales ponderadas proporciona la entrada total o neta de la neurona y, mediante la aplicación de una función matemática - denominada función de salida - , sobre la entrada neta, se calcula un valor de salida, el cual es enviado a otras neuronas.
- Tanto los valores de entrada a la neurona como su salida pueden ser señales excitatorias (cuando el valor es positivo) o inhibitorias (cuando el valor es negativo).



# Modelo de Neurona Artificial

- El elemento básico de computación (modelo de neurona) es un nodo o unidad. Recibe un input desde otras unidades o de una fuente externa de datos. Cada input tiene un peso asociado  $w$ , que se va modificando en el llamado proceso de aprendizaje. Cada unidad aplica una función dada  $f$  de la suma de los inputs ponderadas mediante los pesos. El resultado puede servir como output de otras unidades



$$y_i = \sum_j w_{ij} y_j$$

# Redes Neuronales Artificiales

- Las Redes Neuronales (ANN: Artificial Neural Networks) surgieron originalmente como una simulación abstracta de los sistemas nerviosos biológicos, constituidos por un conjunto de unidades llamadas neuronas conectadas unas con otras.
- El primer modelo de red neuronal fue propuesto en 1943 por McCulloch y Pitts en términos de un modelo computacional de actividad nerviosa. Este modelo era un modelo binario, donde cada neurona tenía un escalón o umbral prefijado, y sirvió de base para los modelos posteriores.



# Características de la Redes Neuronales Artificiales

Las ANN aplicadas están, en general, inspiradas en las redes neuronales biológicas, aunque poseen otras funcionalidades y estructuras de conexión distintas a las vistas desde la perspectiva biológica. Las características principales de las ANN son las siguientes:

- ***Auto-Organización y Adaptabilidad*** : utilizan algoritmos de aprendizaje adaptativo y auto-organización, por lo que ofrecen mejores posibilidades de procesamiento robusto y adaptativo.
- **Procesado no Lineal** : aumenta la capacidad de la red para aproximar funciones, clasificar patrones y aumenta su inmunidad frente al ruido.
- ***Procesado Paralelo***: normalmente se usa un gran número de nodos de procesamiento, con alto nivel de interconectividad.

# Fases de Modelización

- **Fase de entrenamiento:** se usa un conjunto de datos o patrones de entrenamiento para determinar los pesos (parámetros) que definen el modelo de red neuronal. Se calculan de manera iterativa, de acuerdo con los valores de los valores de entrenamiento, con el objeto de minimizar el error cometido entre la salida obtenida por la red neuronal y la salida deseada.
- **Fase de Prueba :** en la fase anterior, el modelo puede que se ajuste demasiado a las particularidades presentes en los patrones de entrenamiento, perdiendo su habilidad de generalizar su aprendizaje a casos nuevos(sobreajuste).



- Para evitar el problema del sobreajuste, es aconsejable utilizar un segundo grupo de datos diferentes a los de entrenamiento, el grupo de validación, que permita controlar el proceso de aprendizaje.
- Normalmente, los pesos óptimos se obtienen optimizando (minimizando) alguna función de energía.
- Un criterio muy utilizado en el llamado entrenamiento supervisado, que consiste en minimizar el error cuadrático medio entre el valor de salida y el valor real esperado.

# Función de Activación

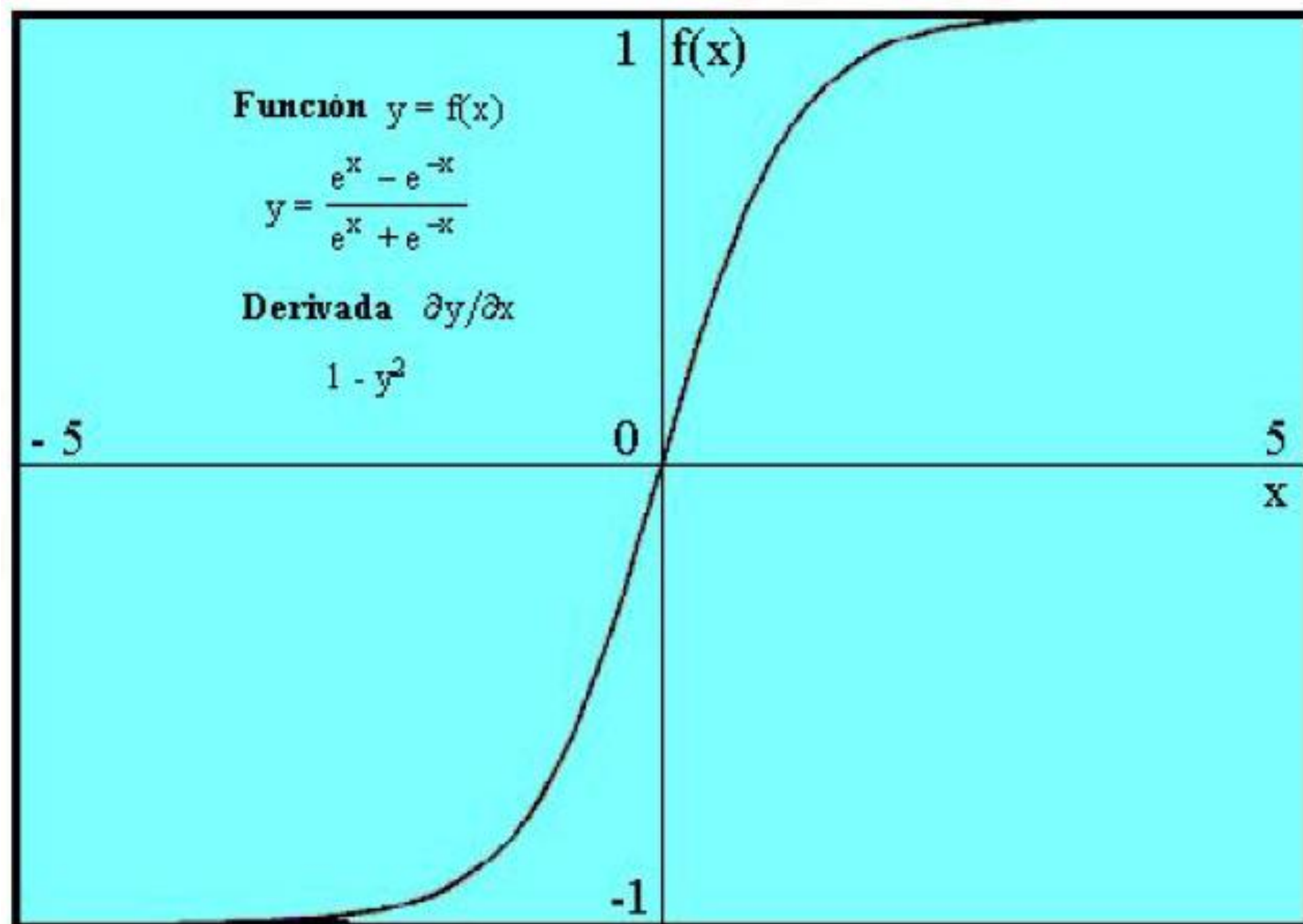
- El valor de red, expresado por la función de base,  $u(w; x)$ , se transforma mediante una función de activación no lineal. Las funciones de activación más comunes son la función logística y la función tangente hiperbólica

**Función Logística:**

$$f(x) = \frac{1}{1 + e^{-x}}$$

**Función hiperbólica:**

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



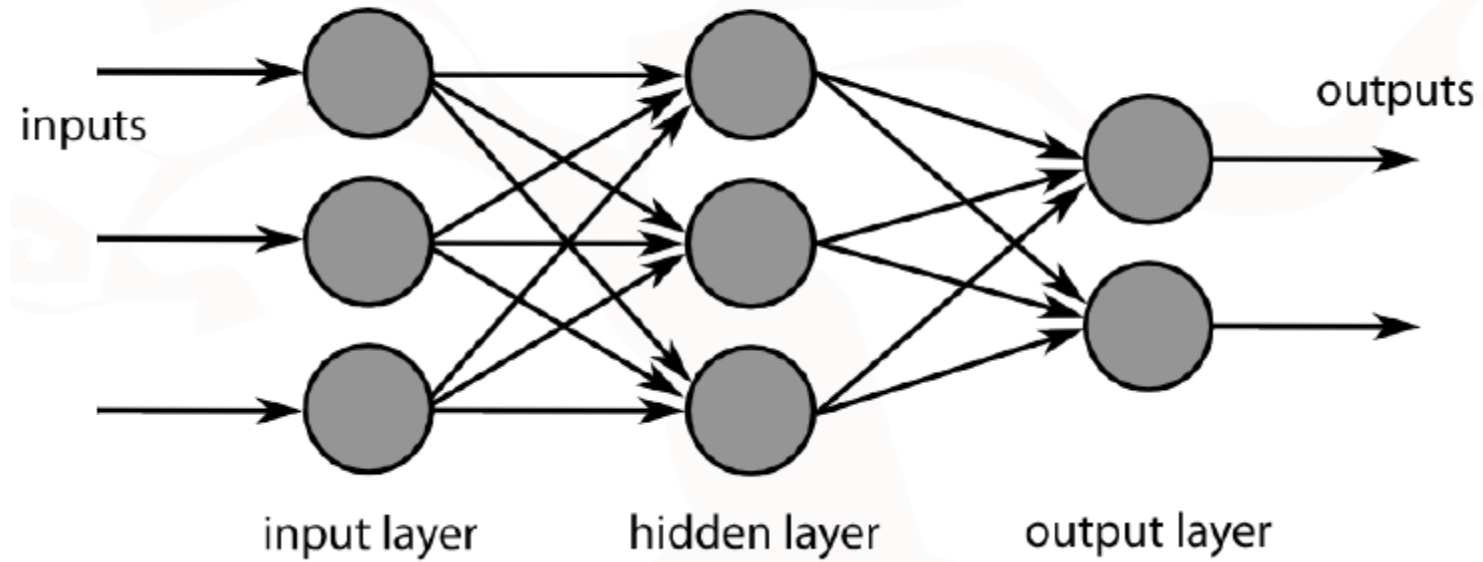
# Estructuras de Conexión

- Una red neuronal se determina por las neuronas y la matriz de pesos. Se pueden definir tres tipos de capas de neuronas:
  - la capa de entrada,
  - la capa oculta y
  - la capa de salida.
- Entre dos capas de neuronas existe una red de pesos de conexión, que puede ser de los siguientes tipos:
  - hacia delante, hacia atrás, lateral y de retardo.

# Tamaño de las Redes Neuronales Artificiales

- En una red multicapa de propagación hacia delante, puede haber una o más capas ocultas entre las capas de entrada y salida. El tamaño de las redes depende del número de capas y del número de neuronas ocultas por capa.
- El número de unidades ocultas está directamente relacionado con las capacidades de la red. Para que el comportamiento de la red sea correcto, se tiene que determinar apropiadamente el número de neuronas de la capa oculta.

# Capas Ocultas



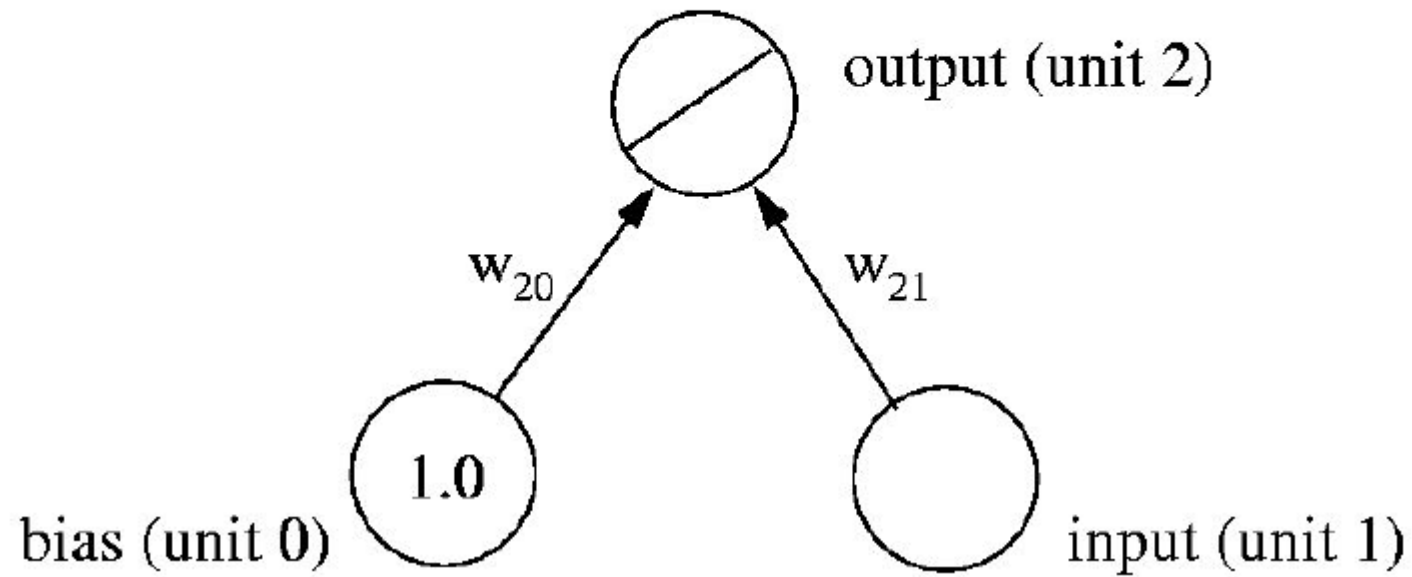


# La Regresión y las Redes Neuronales Artificiales

- Los modelos de regresión estudian la relación entre una serie de variables, denominadas independientes  $x_i$ , y otras variables dependientes o respuesta que se denotan como  $y$ : El objetivo es predecir los valores de  $y$  en función de los valores de  $x_i$ .
- Un modelo lineal se puede implementar como una red neuronal simple: tiene una unidad de *sesgo*, una unidad de *input* y una unidad de *output*. El *input* se refiere a una variable  $x$ , mientras que el *sesgo* siempre es una constante igual a 1.

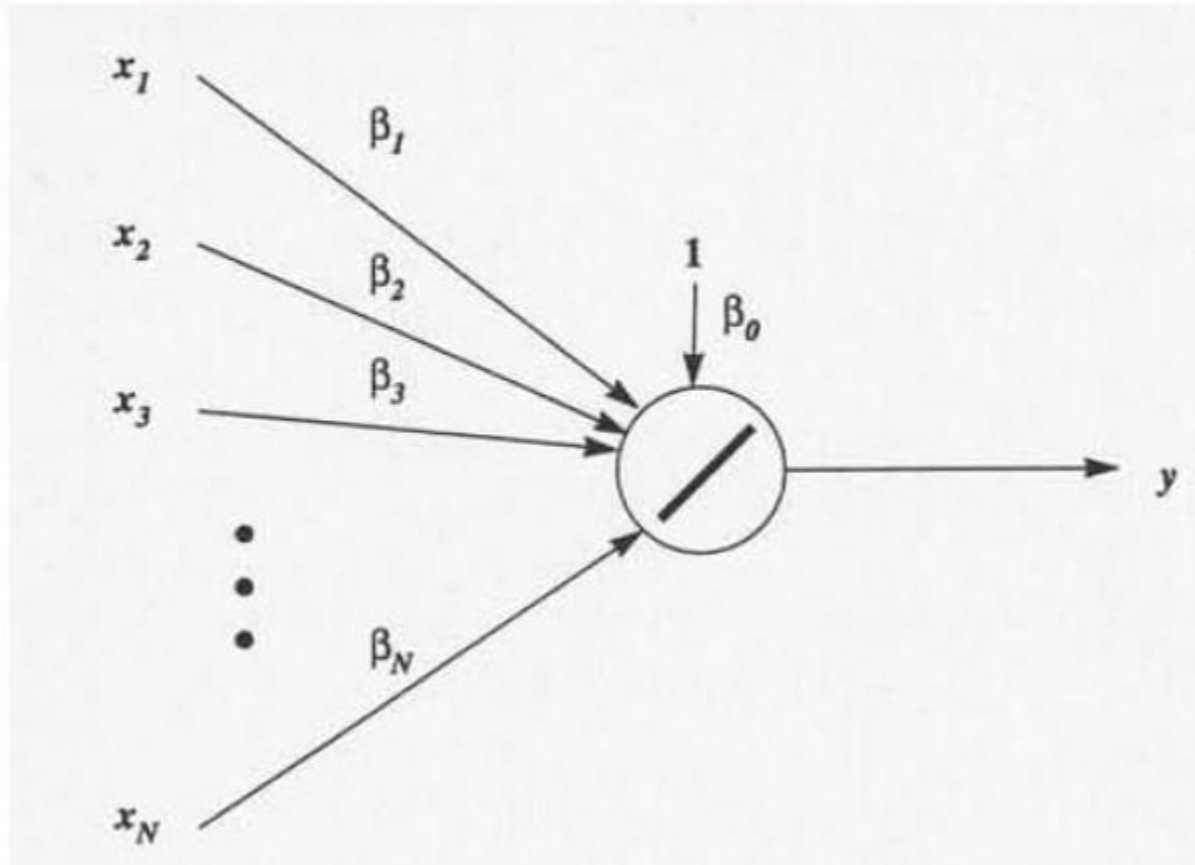
- El *output* podría expresarse pues como:

$$y_2 = y_1 w_{21} + 1.0 w_{20}$$



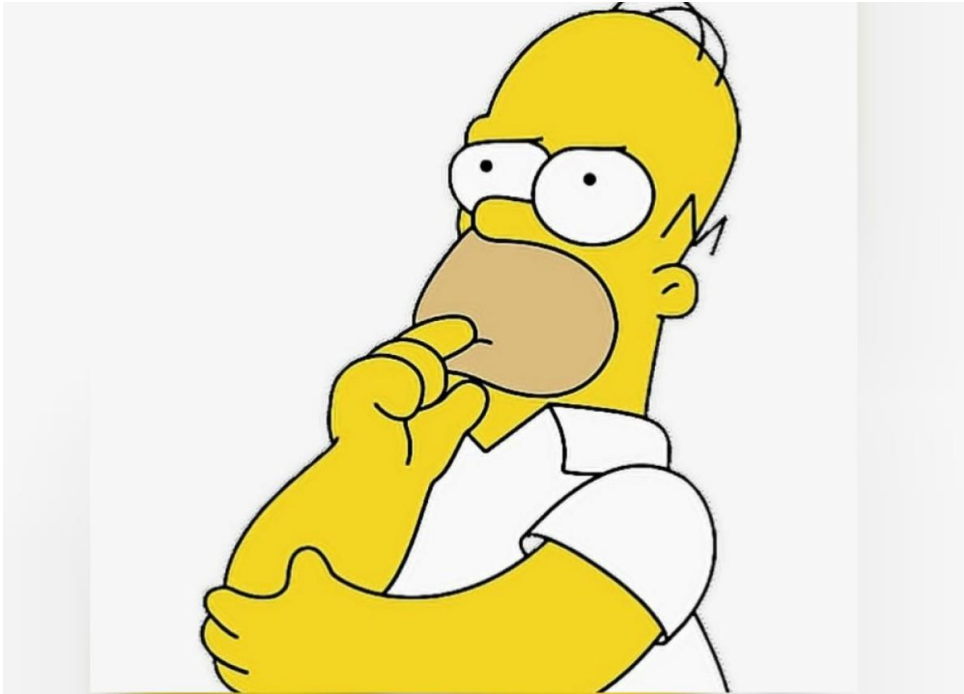
# Modelo Lineal y las Redes Neuronales Artificiales

$$y_p = \sum_{i=1}^N \beta_i x_{pi} + \epsilon$$



# Preguntas

- Alguna pregunta?



# Ejemplo: Redes Neuronales Artificiales

- Utilizar el conjunto de datos de Boston en el paquete MASS.  
Recordemos que el conjunto de datos de Boston es una colección de datos sobre valores de vivienda en los suburbios de Boston. Nuestro objetivo es predecir el valor medio de las viviendas ocupadas por sus propietarios.