

Deep Learning

Física Computacional 2

Ph.D. Santiago Echeverri Arteaga

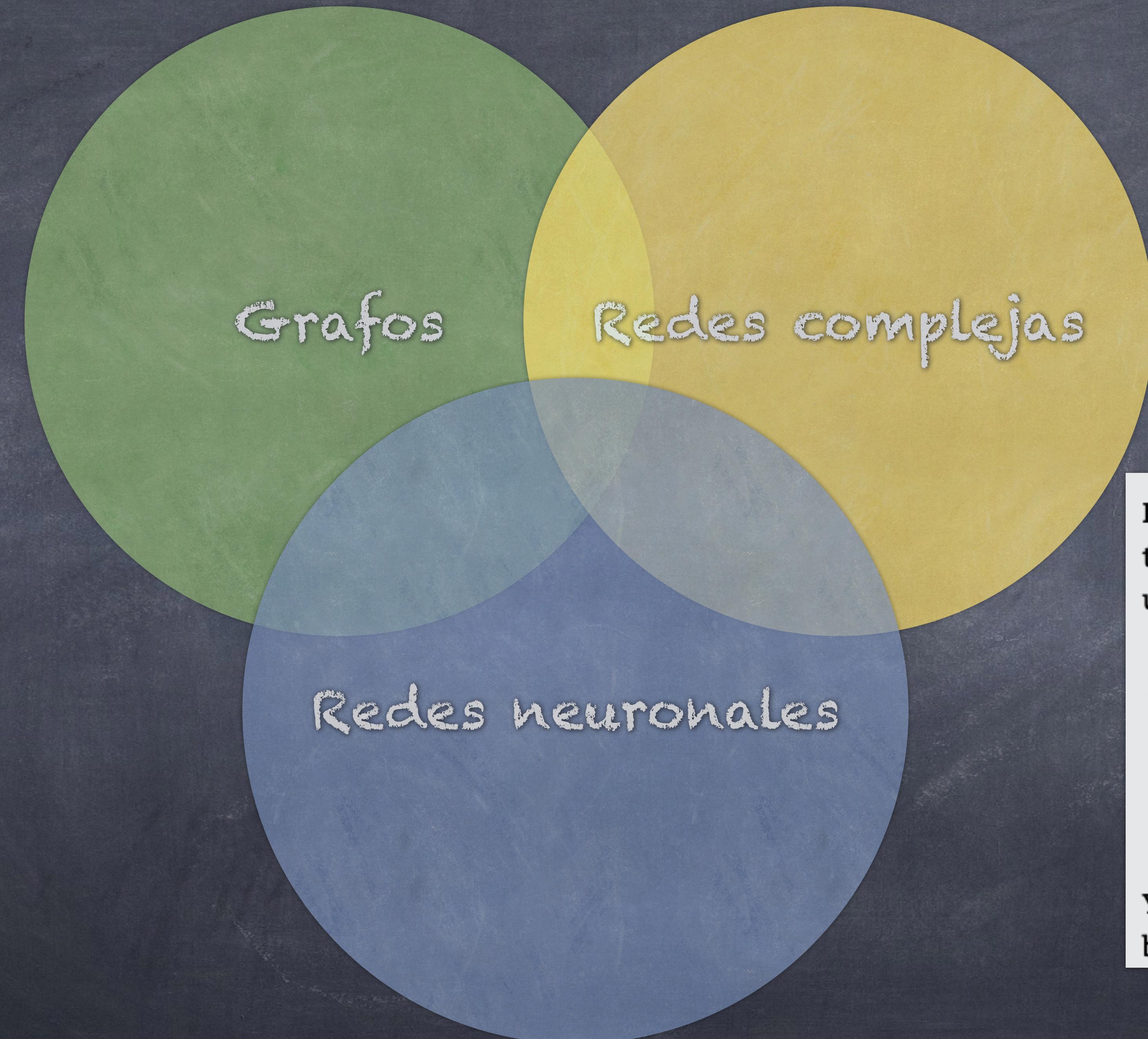




Original
B/N

Original
Color

Reconstrucción
del color



In the scientific literature the terms *network* and *graph* are used interchangeably:

Network Science	Graph Theory
Network	Graph
Node	Vertex
Link	Edge

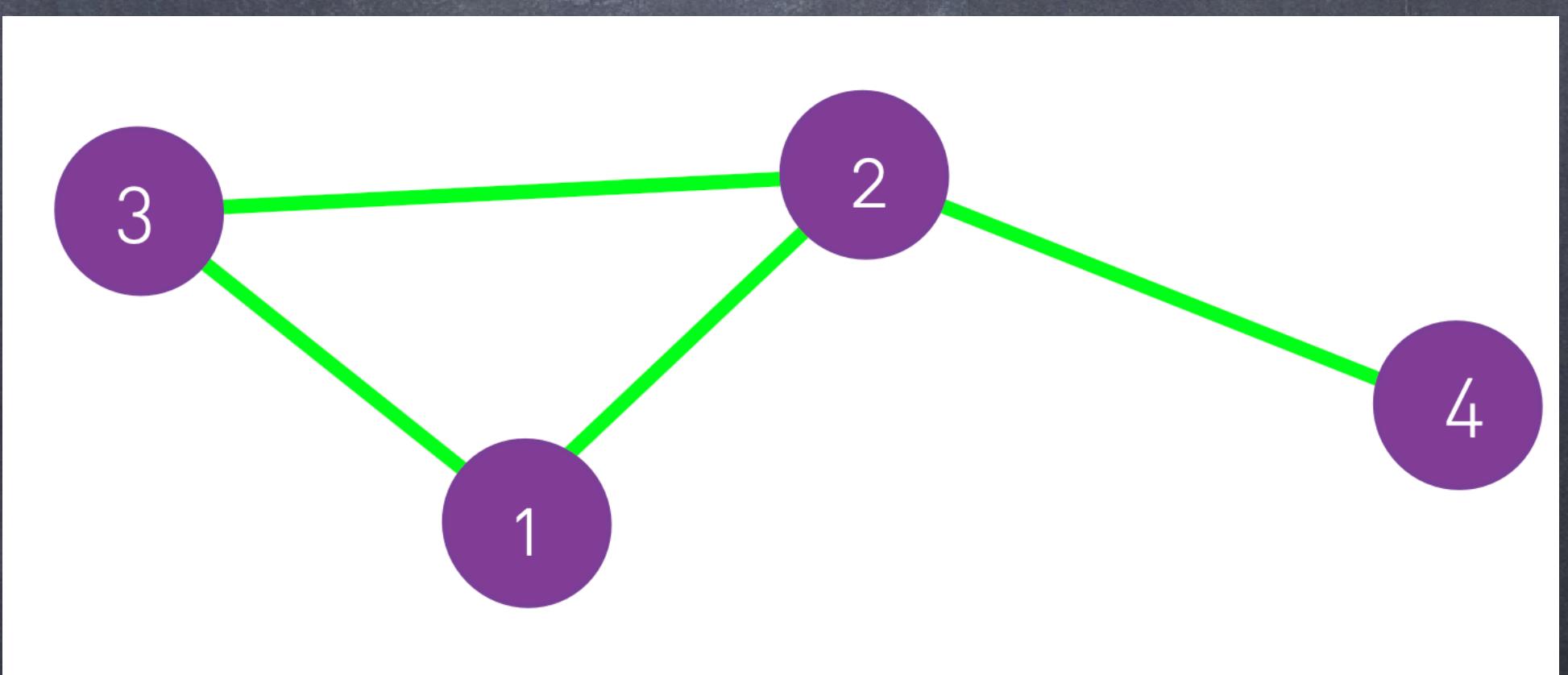
Yet, there is a subtle distinction between the two terminologies:

Definiciones importantes

- Red: "Una red es un catálogo de los componentes de un sistema, a menudo llamados nodos o vértices, y las interacciones directas entre ellos, llamadas enlaces o aristas" Laszlo Barabasi
- Grado del i -ésimo nodo: "El número de enlaces que tiene el nodo en cuestión hacia otros nodos"
- Distribución de grado: "Probabilidad que un nodo seleccionado de forma arbitraria tenga grado K "
- Matriz de adyacencia:
- Matriz de pesos:

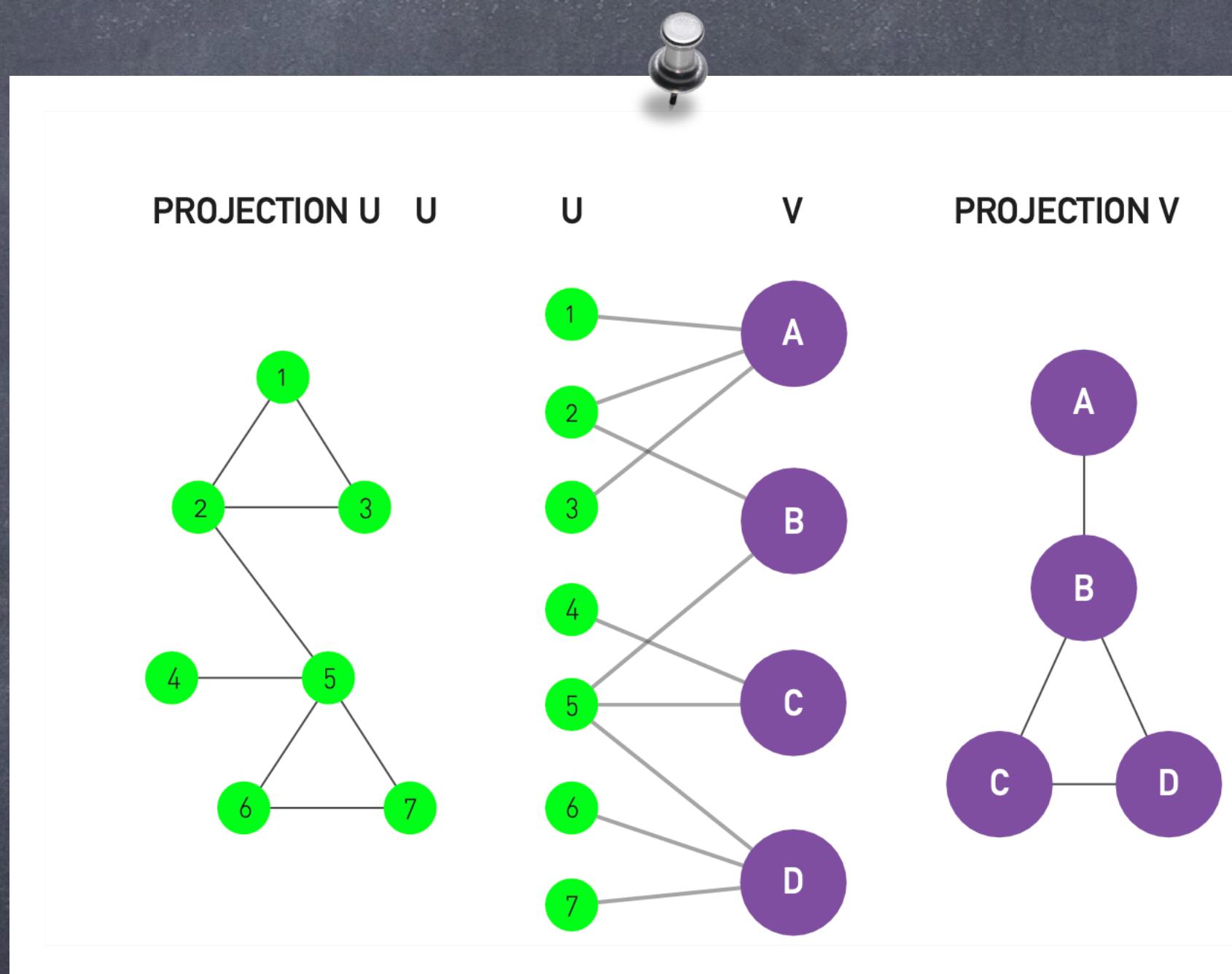
Matriz de adyacencia

- Los enlaces entre los nodos son: $\{(1,2), (1,3), (2,3), (2,4)\}$. Esto puede ser representado por una matriz $N \times N$ (con N nodos) en donde los elementos $A_{ij} = w_{ij}$ representan que tan fuerte está unido el nodo i con el j
- Si $A_{ij} = A_{ji}$ se dice que el enlace no es direccional
- $A_{ii} = 0$
- El grado k_i de un nodo sería $k_i = \sum_{j=1}^N A_{ij}$



Redes bipartitas / multipartitas

- Una red bipartita es una red en la que los nodos los puedo dividir en dos conjuntos disjuntos U, V de tal forma que cada enlace conecte un nodo U con un nodo V .
- Se pueden generar dos proyecciones de la red. La primera contiene solo los nodos U . En esta proyección dos nodos están conectados si ellos se enlazan al mismo nodo de V . (La segunda proyección es análoga pero para V)



(a) RECIPES

INGREDIENTS

COMPOUNDS

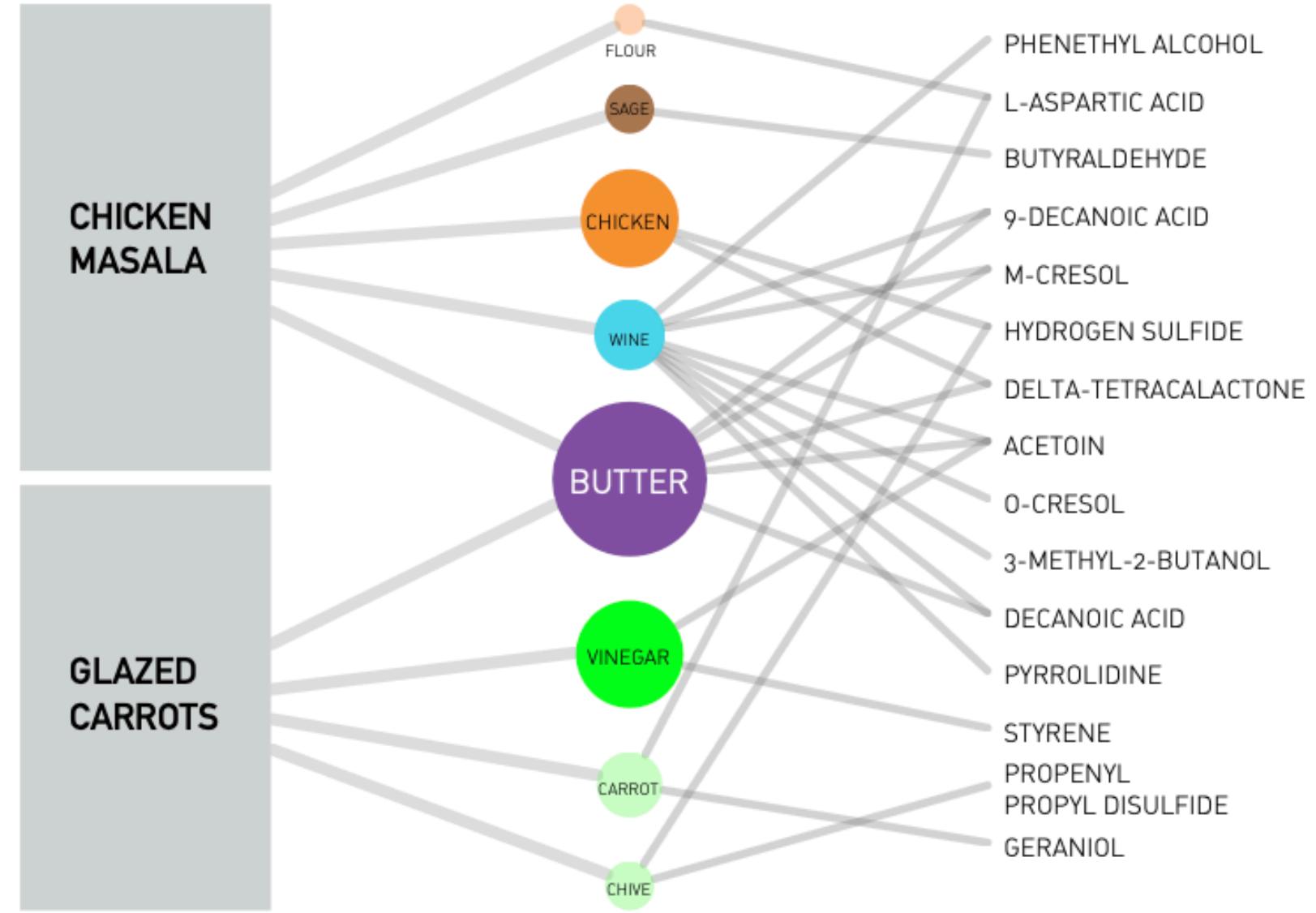
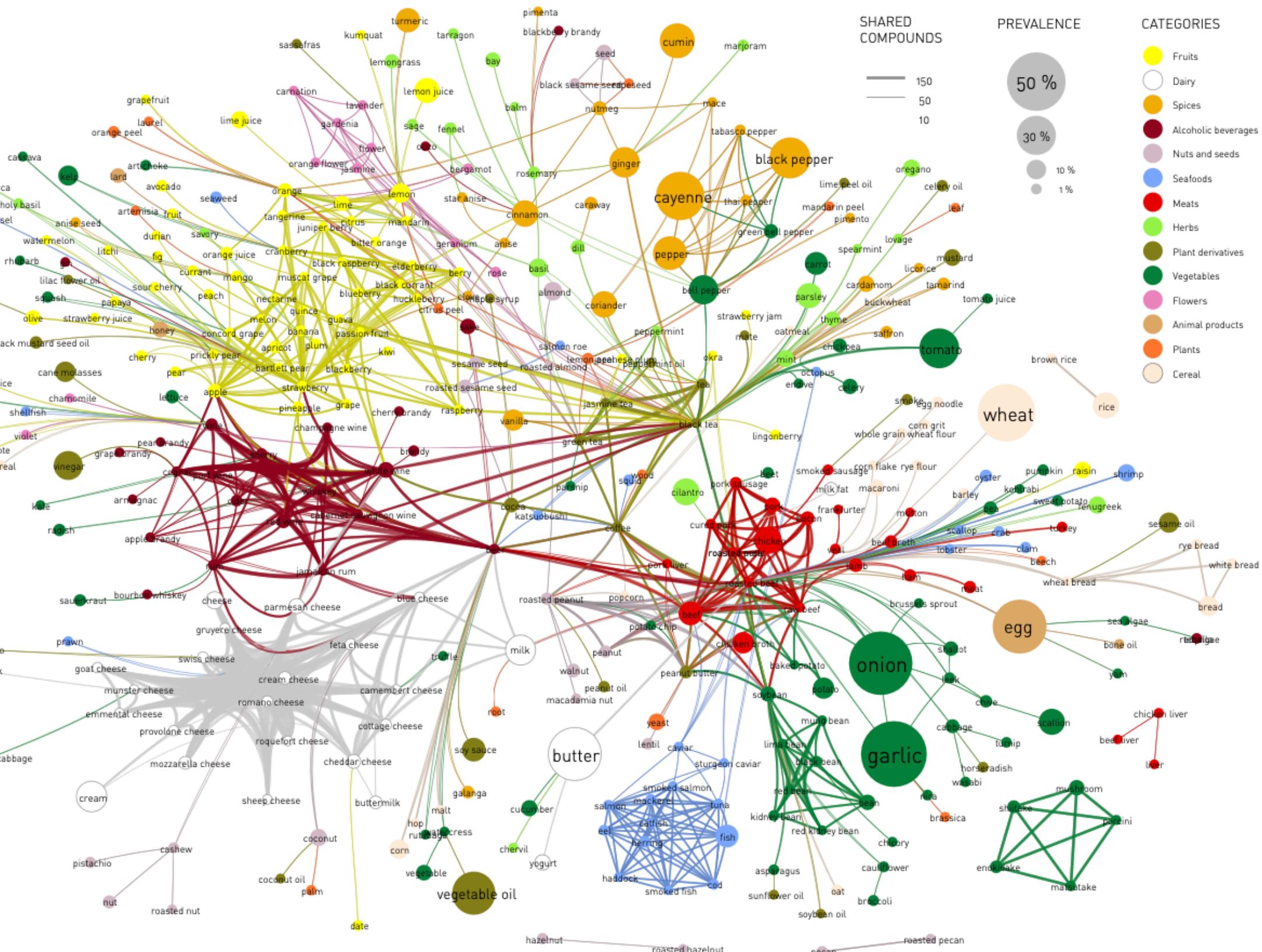


Figure 2.11
Tripartite Network

(a) The construction of the tripartite recipe-ingredient-compound network, in which one set of nodes are recipes, like Chicken Masala; the second set corresponds to the ingredients each recipe has (like flour, sage, chicken, wine, and butter for Chicken Masala); the third set captures the flavor compounds, or chemicals that contribute to the taste of each ingredient.

(b) The *ingredient* or the *flavor network* represents a projection of the tripartite network. Each node denotes an ingredient; the node color indicating the food category and node size indicates the ingredient's prevalence in recipes. Two ingredients are connected if they share a significant number of flavor compounds. Link thickness represents the number of shared compounds.

After [11].

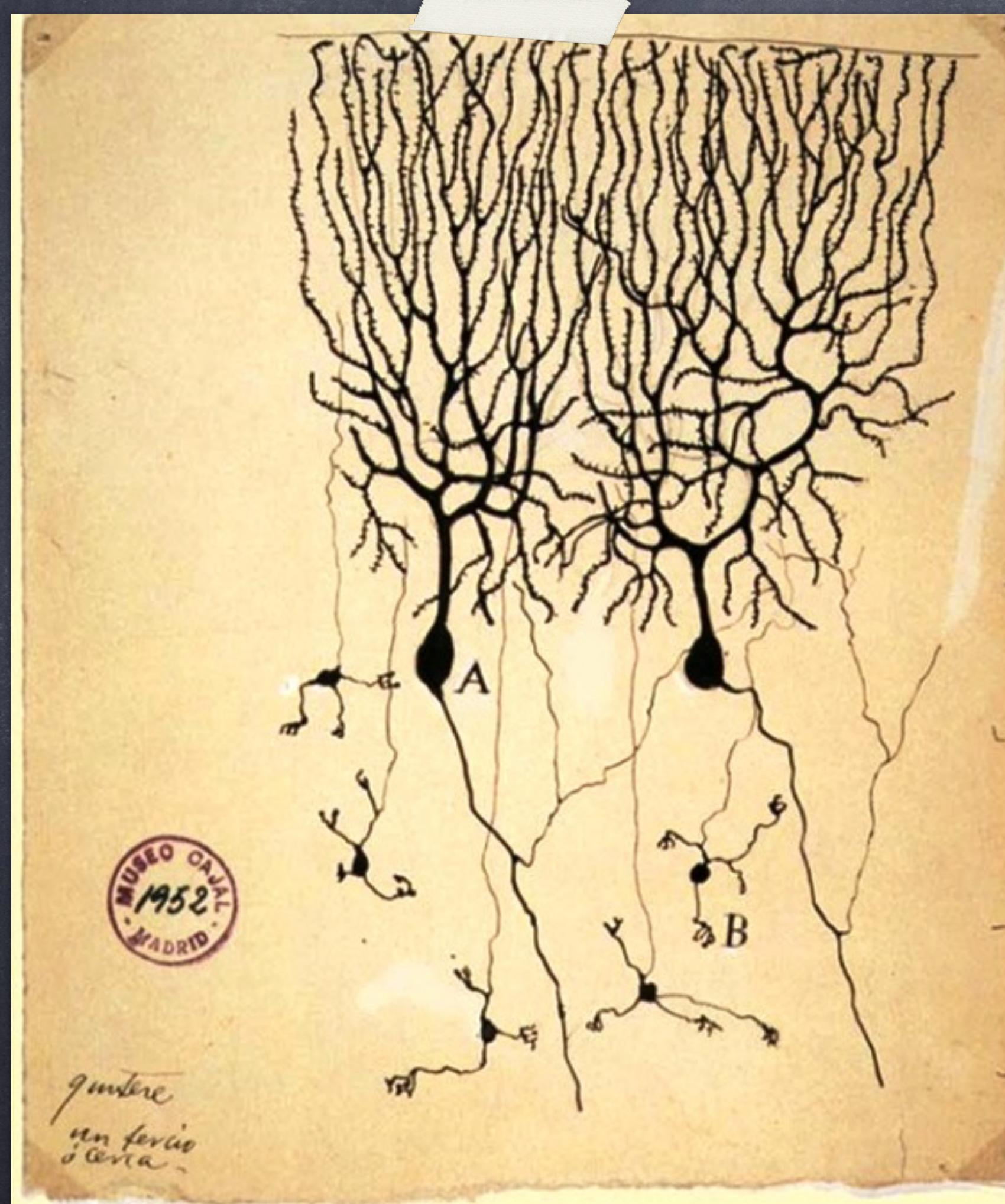


¿Por qué redes
NEURONALES?

PAUS



¿Qué es una neurona?



The Nobel Prize in Physiology or Medicine 1906

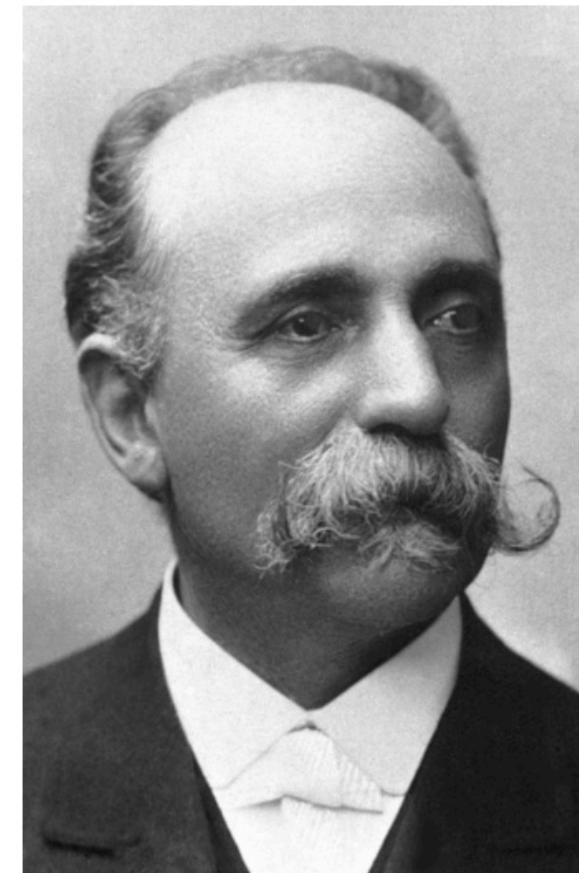
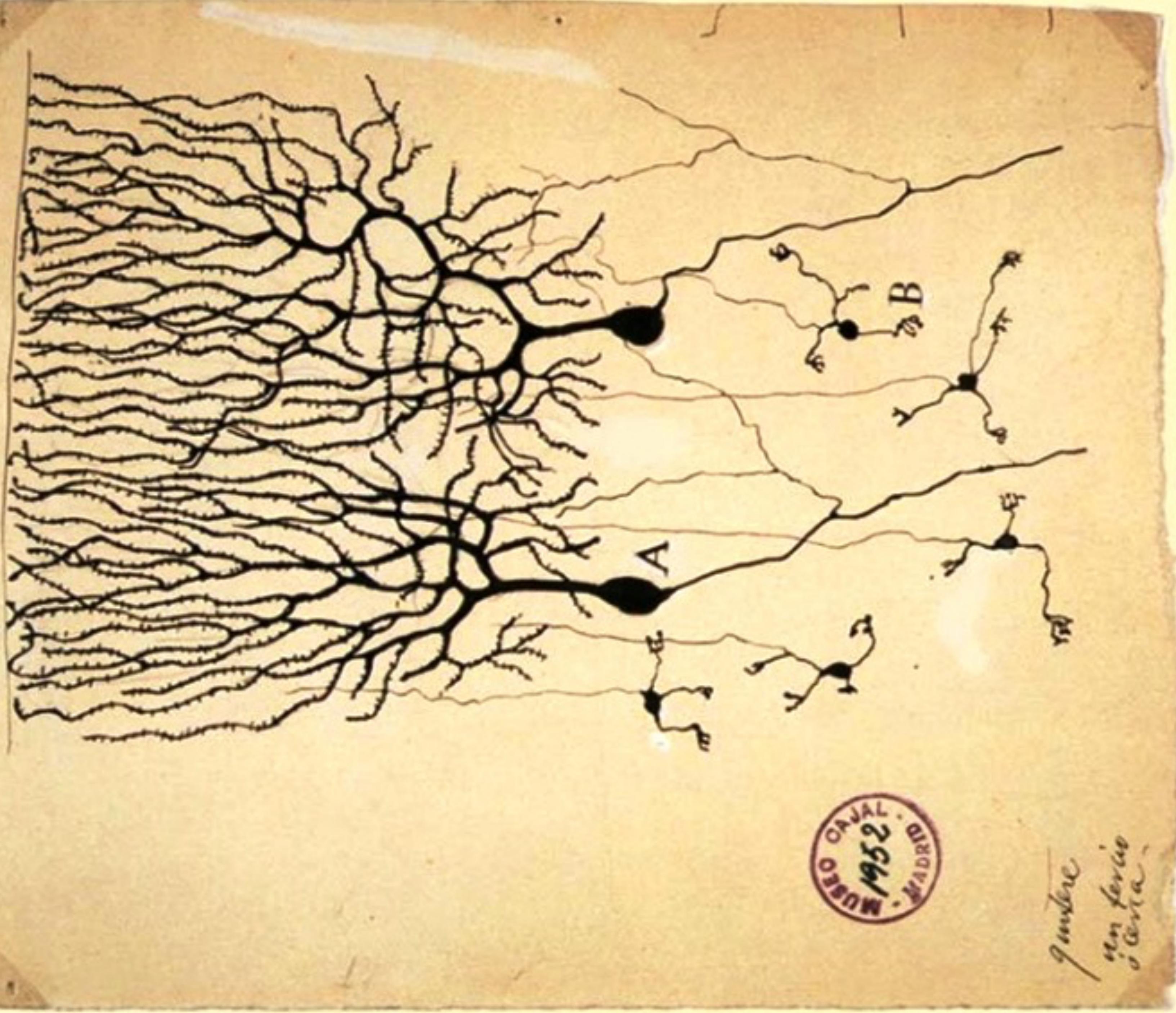


Photo from the Nobel Foundation archive.
Camillo Golgi
Prize share: 1/2

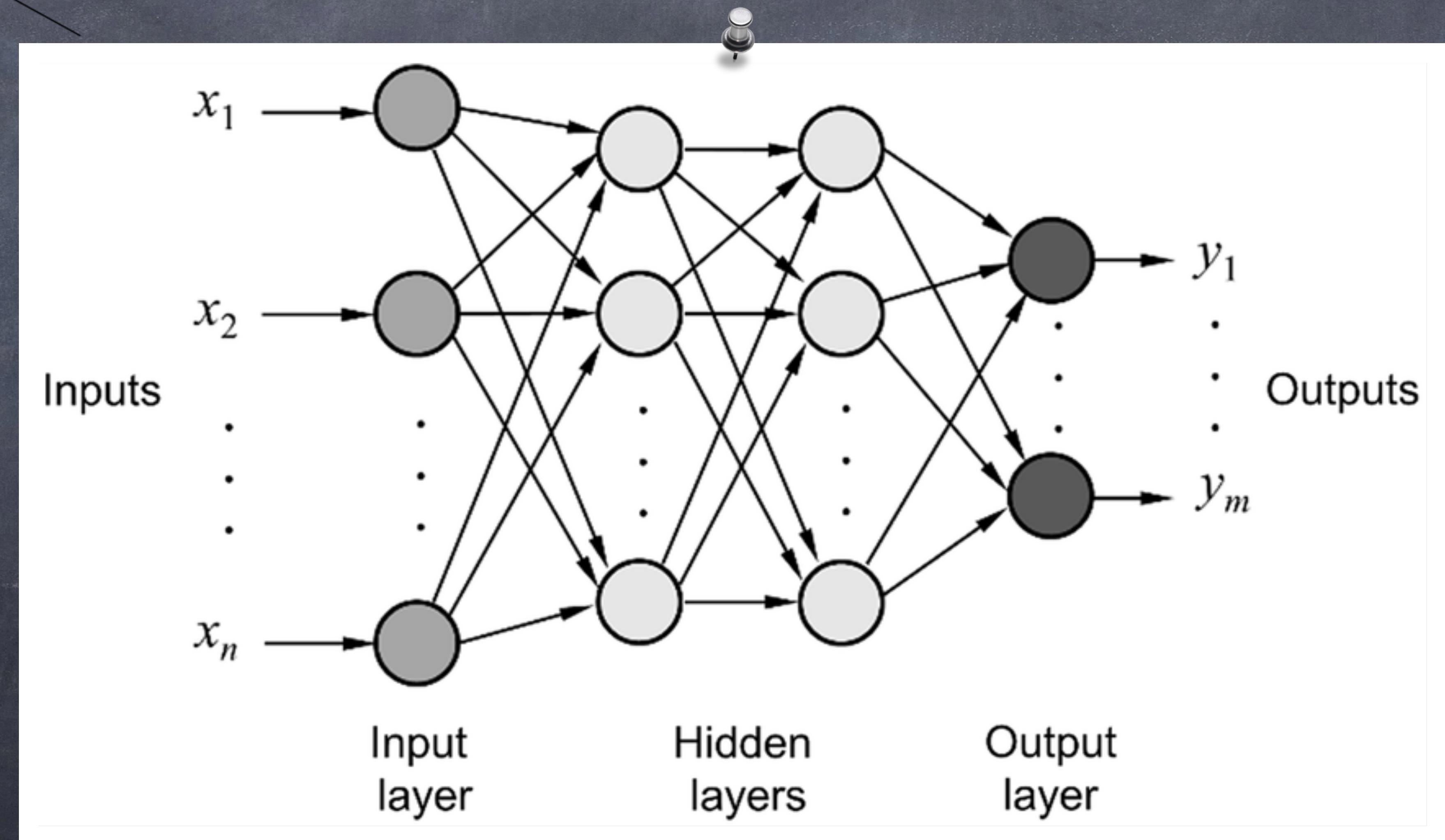
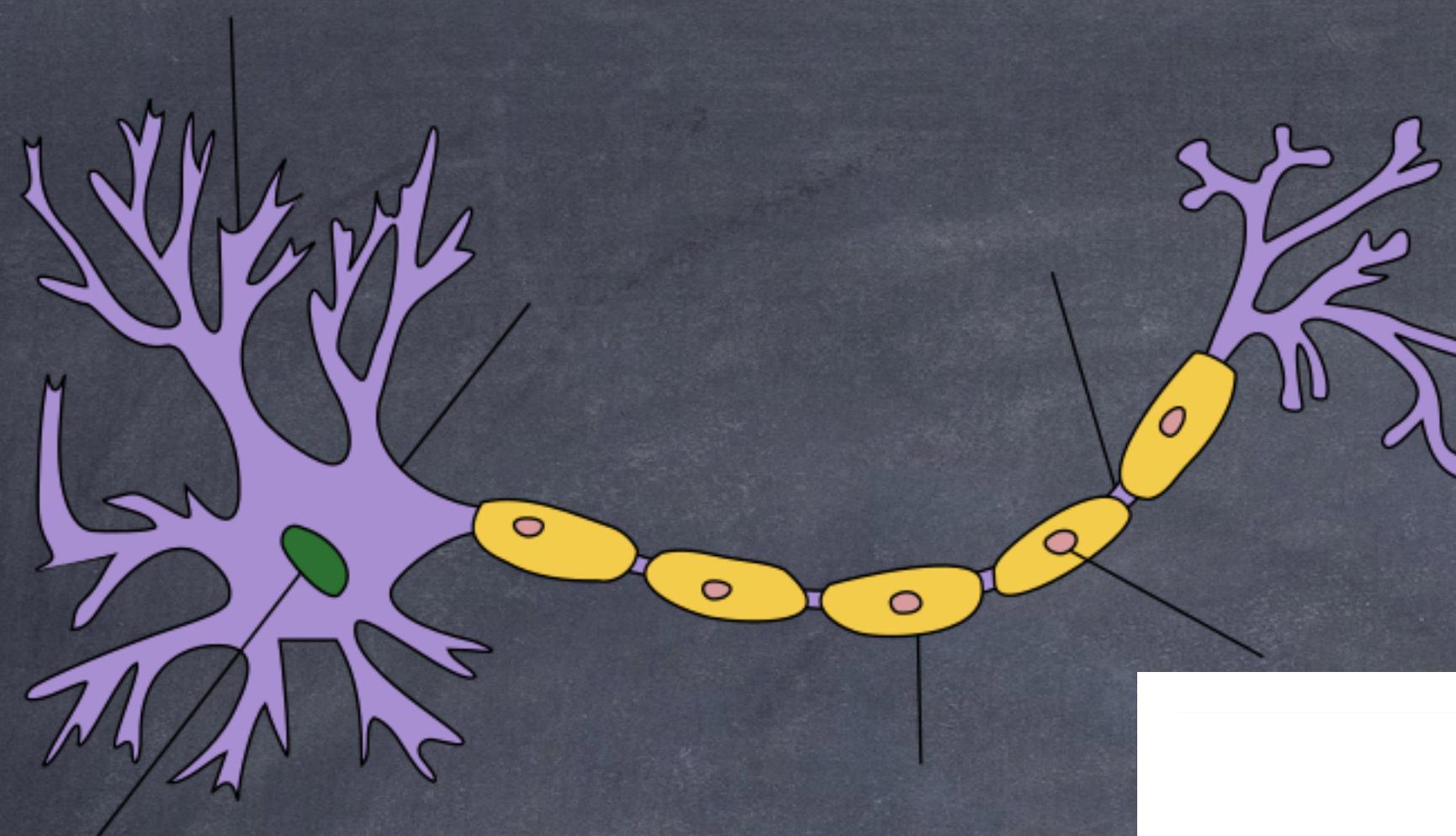


Photo from the Nobel Foundation archive.
Santiago Ramón y Cajal
Prize share: 1/2

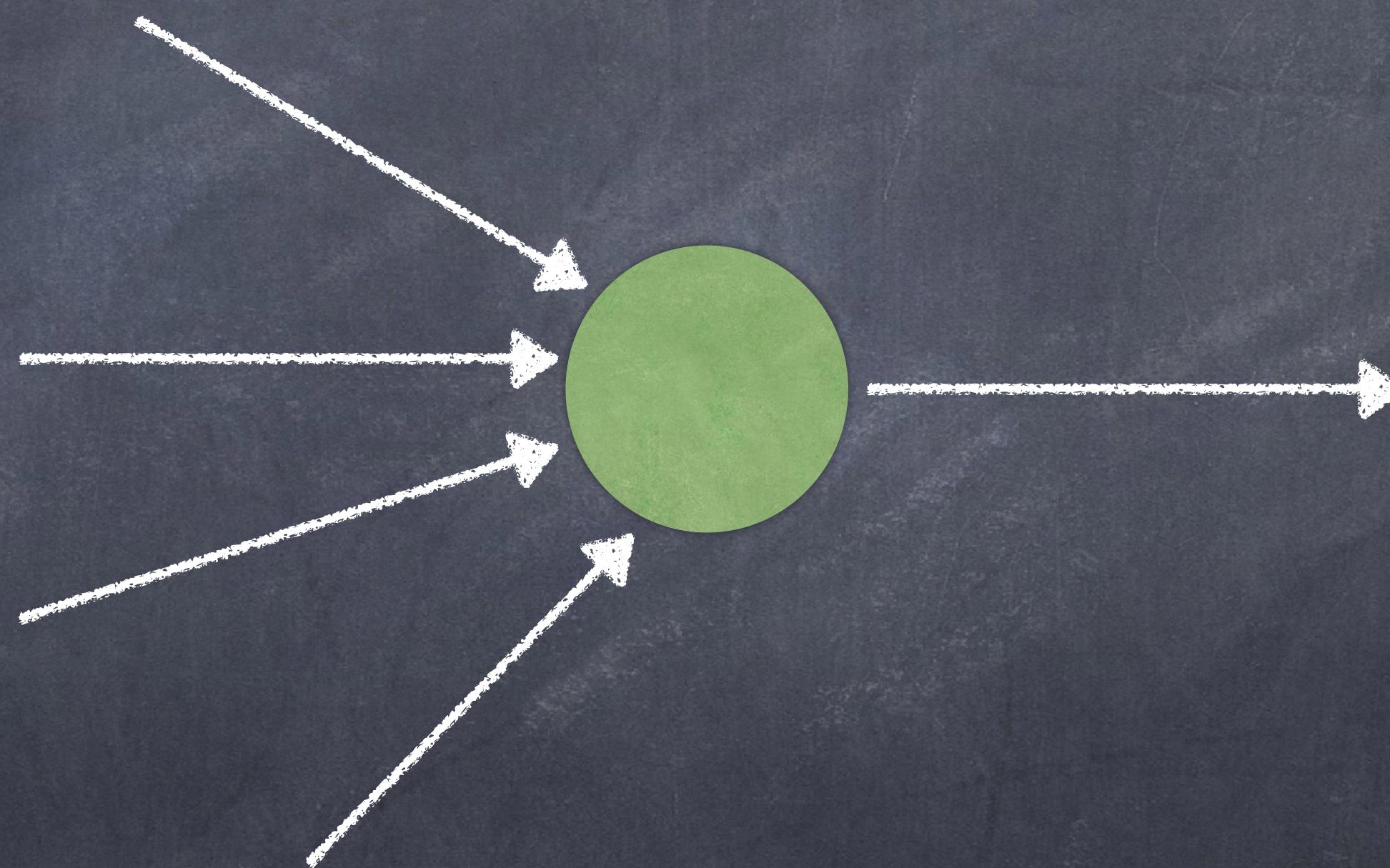
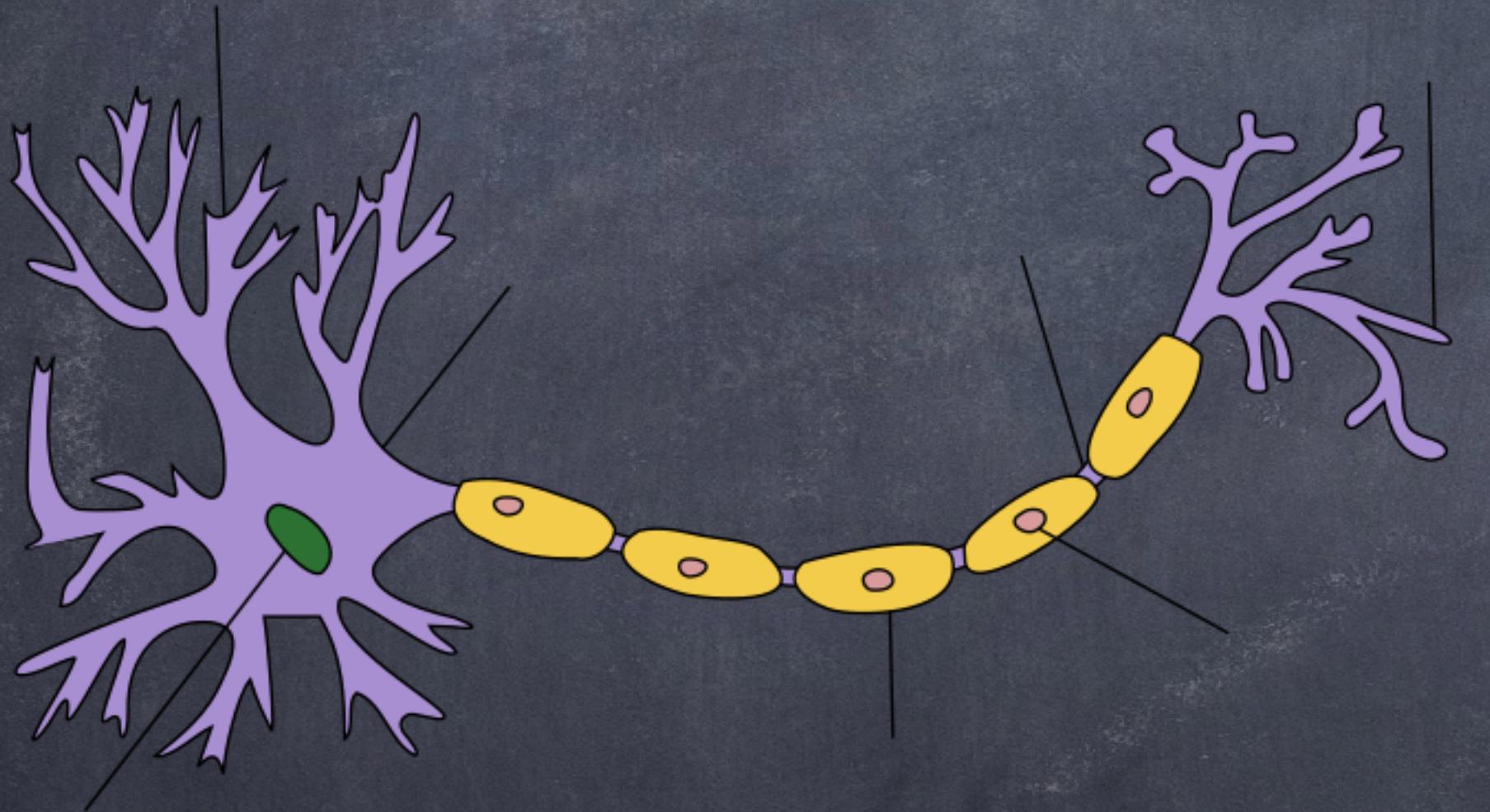
The Nobel Prize in Physiology or Medicine 1906 was awarded jointly to Camillo Golgi and Santiago Ramón y Cajal "in recognition of their work on the structure of the nervous system"



Red
Neuronal



Percepción: El nombre para nuestra neurona (Nodo)



Regresión Lineal VS Regresión Logística: Función Sísmoide

- La regresión logística no es más que una regresión lineal con una función de activación
- Función de activación: Función no lineal que se aplica a cada elemento en la salida de una neurona
- Cuál es la derivada de la función sísmoide

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

¿Por qué una red y no solo un perceptrón?

- Un percepción implica frontera de clasificación lineal
- Vamos a tener una red multipartita con al menos tres subconjuntos. A cada subconjunto lo llamaremos layer
- Input Layer: Características iniciales a ser transferidas a la red. Cada nodo de esa capa corresponde a una observación, que en general es un tensor.
- Hidden Layers: Características intermedias. Cada nodo corresponde a dos transformaciones sucesivas, una transformación lineal y una no lineal que aplica elemento a elemento.
- Output Layers: Características predichas. Cada nodo corresponde a un tensor.
- Hay diferentes arquitecturas de redes. Entre una y otra cambia la conexión entre nodos y el tipo de transformaciones que se hacen
- Hay métodos de regularización que afectan las conexiones de la red.
- Una red neuronal es una forma de representar un campo tensorial no lineal compuesto
- Una red neuronal se llama red neuronal profunda si tiene más de dos capas ocultas

Redes Neuronales

Feedforward
Neural Nets

Convolutional
Neural Nets

Inception

ResNet

Recurrent
Neural Network

LSTM

GRU

Seq2Seq

GAN

Máquina de
Boltzman
Restringida

Autoencoder

Deep Belief
Network

Autoencoder
variacional

Funciones de activación

- **Sigmoide:** $\sigma(z) = \frac{1}{1 + e^{-z}}$

- Entre 0 y 1, la derivada es sencilla, problema de Vanishing Gradients

- **Softmax:** $\sigma : \mathbb{R}^k \rightarrow [0,1]^k$

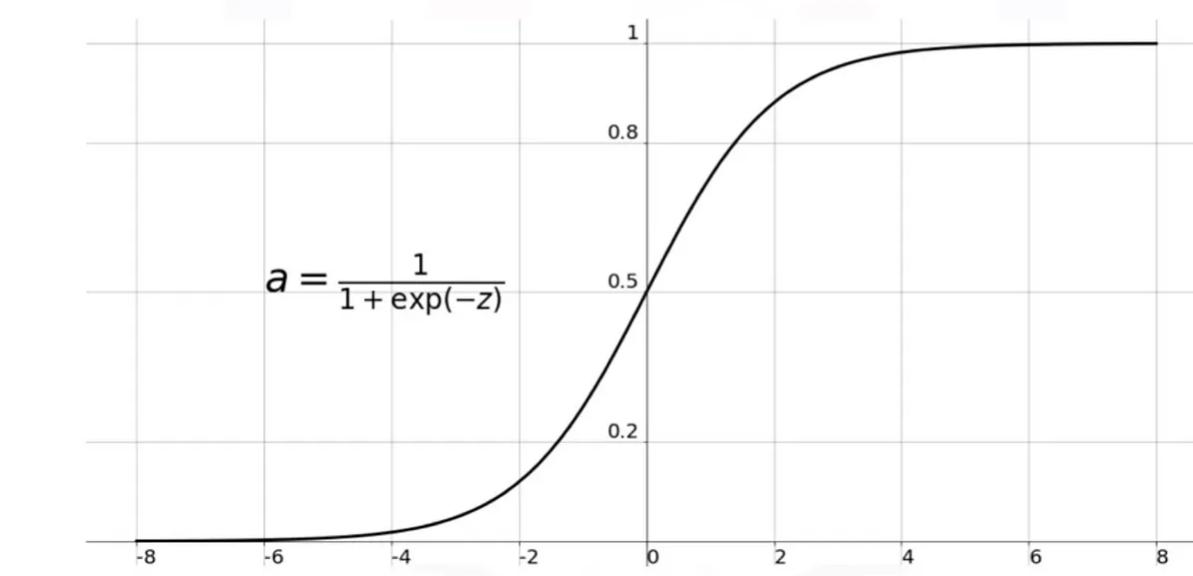
$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{i=1}^k e^{z_i}}$$

- Como Sigmoide pero multiclase

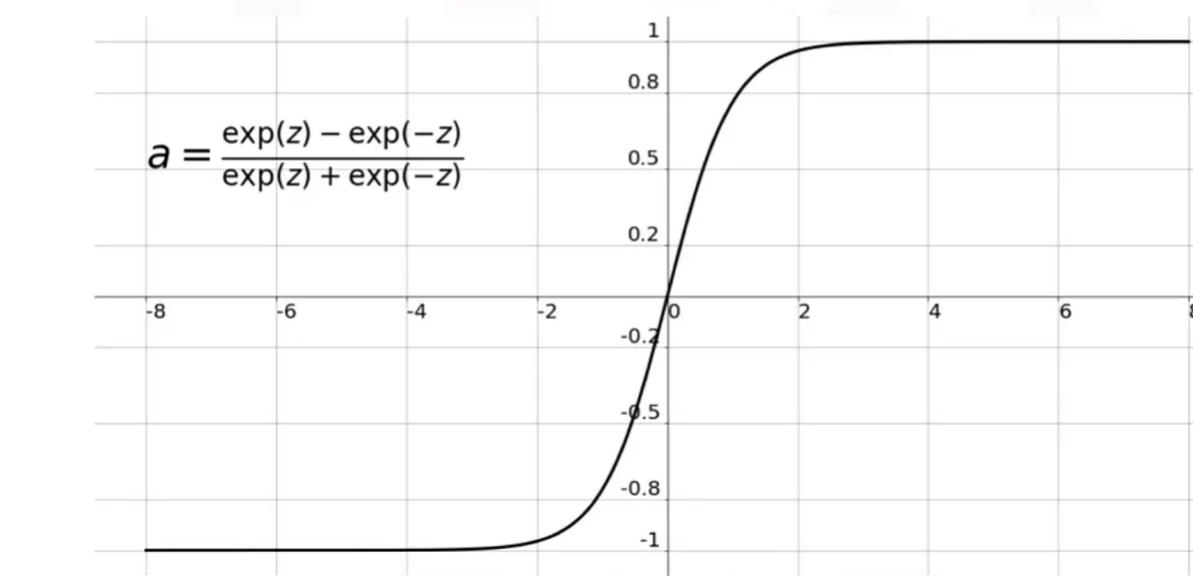
- **Tangente Hiperbólica:** $\tanh(z) = \frac{e^{2x} - 1}{e^{2x} + 1}$

- Entre -1 y 1, Vanishing Gradients cuando los valores $|z| > 2.5$

Sigmoid Function



Hyperbolic Tangent Function



Softmax Function

$$z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1.6 \\ 0.55 \\ 0.98 \end{bmatrix}$$

$$a_i = \frac{e^{z_i}}{\sum_{k=1}^m e^{z_k}}$$

↓
Softmax

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.51 \\ 0.18 \\ 0.31 \end{bmatrix}$$

Funciones de activación

- Rectified Linear Unit ReLU:
$$\begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$
- Útil para efectos grandes, hace cero algunos nodos, no tiene Vanishing Gradients, eficiente y mantiene información de cada nodo
- Leaky ReLU:
$$\begin{cases} az & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases}$$
- Como ReLU pero sin hacer cero los nodos

