

Artificial Neural Networks

Martín Gutiérrez

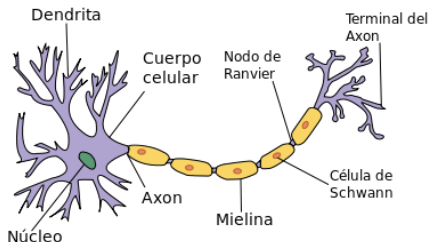
August 7, 2022

YAY!!! ANNs... La llean!!!

Este tema abarca varias sesiones, pues hay varios tipos de ANNs, aunque el concepto de base siempre es el mismo.

De dónde vienen?

Vienen del concepto de una neurona del cerebro, por ende, echemos un vistazo a la estructura de la neurona biológica.



Una neurona artificial

La neurona biológica se basa en emisión y recepción de impulsos químicos/eléctricos. Así entonces, se puede abstraer una neurona a un elemento que consume una entrada (impulso químico o eléctrico) y que emite una salida (otro impulso químico o eléctrico). Es (típicamente) unidireccional.

Así entonces, en similitud a una operación, se podría asociar a una función matemática en la que hay entradas y genera una o varias salidas.

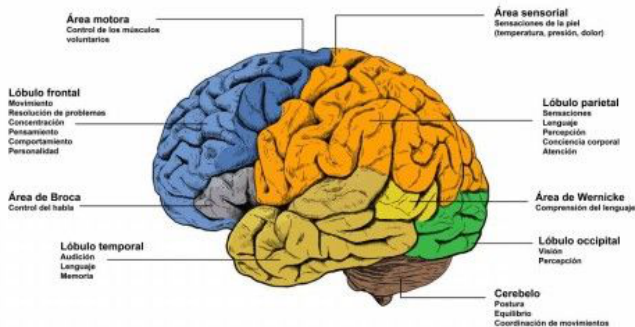
Redes de neuronas (I)

La función del cerebro está principalmente sustentada en un proceso de transferencia (por relevos) de los impulsos denominado sinápsis. Las sinápsis son instancias de contacto entre neuronas a través de los impulsos. Los impulsos llegan y son “procesados” a través de una secuencia de operaciones paralelas (de transferencia).

En realidad, es la morfología de esas conexiones - sinápsis - las que determinan la función, puesto que dicha forma es la que dirige el flujo de impulsos a través de las neuronas. Así pues, es el flujo de impulsos descrito el que inspira a las Artificial Neural Networks (ANN).

Redes de neuronas (II)

Existen otras características morfológicas que también inspiran a las ANN: la regionalidad de la función.



Qué sucede en circunstancias de daño a una región? En qué consiste la rehabilitación?

Establecimos que la inspiración de los elementos base de las ANN son las neuronas biológicas. Además, se vinculan esas neuronas a neuronas artificiales, que indicamos anteriormente, corresponden a operaciones matemáticas.

Específicamente, es una función que opera sobre todas las entradas que le llegan y entrega una salida que es relevada a todas aquellas neuronas que se conectan con ella.

Perceptrón (II)

Formalmente, un perceptrón es una neurona artificial que se basa en el evento de “gatillado”: pondera las entradas (y el bias) por ciertos “pesos”, los agrega y evalúa (gracias a una función de activación) si debe enviar un impulso.

Esto se parece a cierta función vista con anterioridad...

$$y = \theta^T X$$

La diferencia radica en que se pueden conectar los elementos con profundidad arbitraria, generando así un perceptrón multicapa, que ya es la primera forma de ANN (feed-forward).

Funciones de activación (I)

Recordemos la función step (umbral) y a la vez integremos con la expresión mostrada:

$$y = \text{step}(\theta^T X) = \begin{cases} 1 & \text{si } \theta^T X > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Se ve similar a algo conocido?

A la expresión de y se le conoce como función de activación. Controla el relevo del impulso hacia adelante de la ANN en cada neurona.

Funciones de activación (II)

Hay otras opciones...

- **Gaussiana:** $f(X) = e^{-\frac{1}{2\sigma} \sum_{i=1}^n (x_i - c_i)^2}$, con $\sigma, c_1, c_2, \dots, c_n$ constantes y la salida de la función es dependiente de la distancia entre X y C .
- **Sigmoidal:** $f(X) = \frac{1}{1 + e^{-\sum_{i=1}^n (\theta_i x_i)}}$. Esta función es una alternativa no discreta y no lineal a la función umbral. Se le conoce también como función *logística*. La ventaja de esta función por sobre la función umbral es que es continua y diferenciable.
- **Tangente Hiperbólica:** $f(X) = \frac{e^X - e^{-X}}{e^X + e^{-X}}$. Es una función similar a la función *logística* con la salvedad que existe dentro del rango $[-1, 1]$.

Funciones de activación (III)

Otras opciones más...

- **ReLU:** $f(X) = \begin{cases} 0 & \text{si } X < 0 \\ X & \text{en otro caso} \end{cases}$. Es una función típicamente destinada a predicción y presenta una alternativa a las funciones mencionadas en la slide anterior.
- **Softmax:** $f(X) = \frac{e^{x_i}}{\sum_{j=1}^K (e^{x_j})}$. Es una función que entrega una importancia relativa a las clases presentadas. La suma de todos los valores entregados es de 1.

Aprendizaje supervisado? o no supervisado?

Depende. Hay muchas formas distintas de ANN y dependiendo de cada tipo es si opera bajo aprendizaje supervisado o no.

Sin embargo, muchas veces nos estaremos refiriendo a ese tipo de ANN y por ende, la gran mayoría de los casos, se estará en presencia de una herramienta para aprendizaje supervisado.

Clasificación? o regresión?

Depende. Puede cumplir con ambas tareas. Naturalmente, hay ciertos tipos de ANN que se ajustan mejor a clasificación que a regresión y vice-versa.

Típicamente, las redes feed-forward son capaces de llevar a cabo cualquiera de las tareas indistintamente.

En esto, las ANN se asemejan mucho a otros algoritmos de aprendizaje de máquina como regresión lineal/logística o SVMs (específicamente, en qué aspecto?).

La relación es la misma que hay entre regresión lineal y regresión logística.

Además, la separabilidad lineal explica la incapacidad de un perceptrón de aprender XOR.

Aún habiendo hablado de generalidades de ANN, falta mucho por recorrer, por ejemplo: detalles del algoritmo de aprendizaje, distintos tipos de ANNs, overfitting, etc.