



Aprendizaje competitivo

Redes Neuronales, DC-FCEyN-UBA

Rosana Matuk
Primer Cuatrimestre 2017

- ☐ *Aprendizaje No Supervisado*
- ☐ *Aprendizaje Competitivo*
- ☐ *Aplicación: Vector Quantization*

Aprendizaje no supervisado

El aprendizaje competitivo se enmarca en el paradigma del aprendizaje no supervisado.

El aprendizaje no supervisado no requiere un teacher. Durante el entrenamiento, la red neuronal recibe un número de patrones de entrada, y descubre características o dependencias de los datos. El aprendizaje no supervisado tiende a seguir la organización neuro-biológica del cerebro.

Aprendizaje no supervisado: importancia



Yann LeCun

13 de marzo de 2016 · 🌐

✓ Siguiendo ▼



Statement from a Slashdot post about the AlphaGo victory: "We know now that we don't need any big new breakthroughs to get to true AI". That is completely, utterly, ridiculously wrong.

As I've said in previous statements: **most of human and animal learning is unsupervised learning. If intelligence was a cake, unsupervised learning would be the cake**, supervised learning would be the icing on the cake, and reinforcement learning would be the cherry on the cake. We know how to make the icing and the cherry, but we don't know how to make the cake.

We need to solve the unsupervised learning problem before we can even think of getting to true AI.

¿Cuánta información necesita predecir el sistema?

■ Reinforcement Learning (cherry)

- ▶ The machine predicts a scalar reward given once in a while.
- ▶ **A few bits for some samples**

■ Supervised Learning (icing)

- ▶ The machine predicts a category or a few numbers for each input
- ▶ **10→10,000 bits per sample**

■ Unsupervised Learning (cake)

- ▶ The machine predicts any part of its input for any observed part.
- ▶ Predicts future frames in videos
- ▶ **Millions of bits per sample**



Aprendizaje no supervisado: importancia

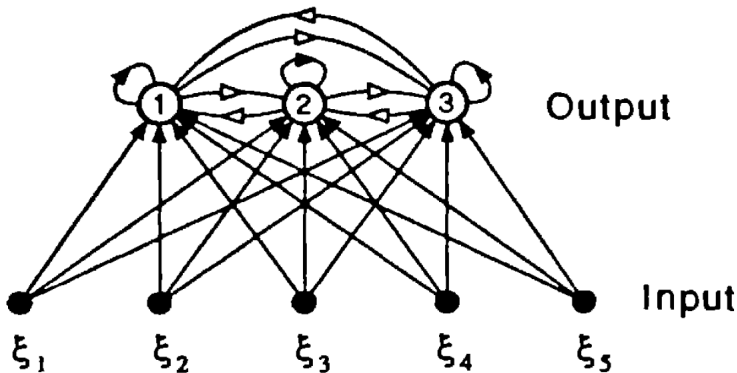


- ☒ *Aprendizaje No Supervisado*
- ☐ *Aprendizaje Competitivo*
- ☐ *Aplicación: Vector Quantization*

Las redes neuronales competitivas usualmente aprenden **matcheando cierto criterio de similaridad**. Típicamente, las redes competitivas sirven para aprender propiedades estadísticas de los inputs.

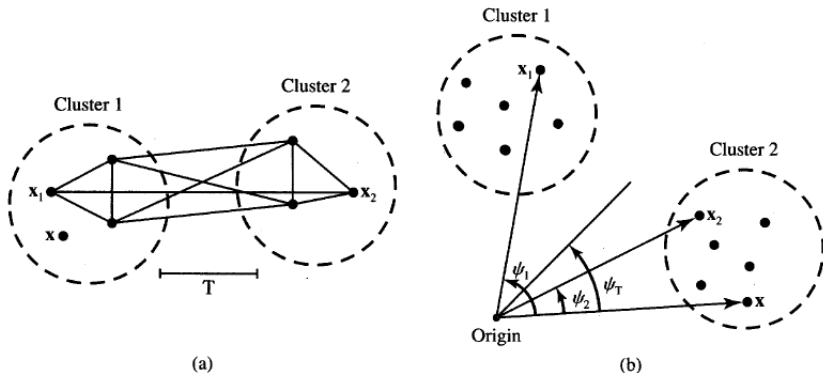
En una red neuronal competitiva, las neuronas compiten entre ellas para determinar la neurona de salida que disparará. **La neurona más similar al patrón de entrada será la neurona ganadora.**

Aprendizaje Competitivo: Arquitectura



Red competitiva: Las flechas blancas son inhibitorias y las flechas negras con excitatorias. La neurona ganadora inhibe a todas las otras.

Algunas Medidas de Similaridad



(a) Distancia euclídea: $\|x - x_i\| = \sqrt{(x - x_i)^t (x - x_i)}$

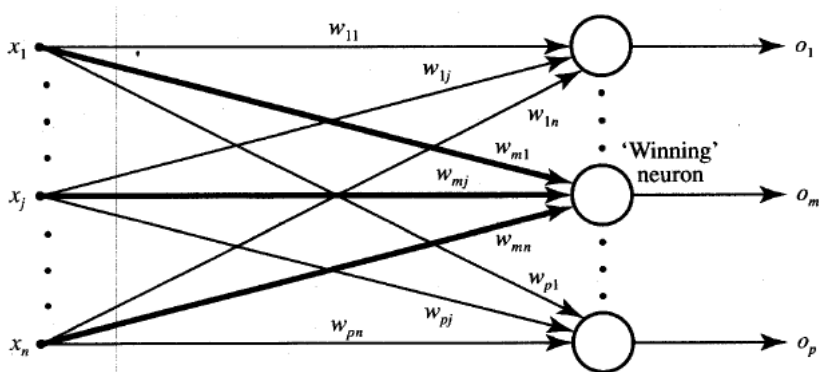
(b) Producto escalar normalizado: $\cos \psi = \frac{x^t x_i}{\|x\| \cdot \|x_i\|}$

Algunas Medidas de Similaridad: observaciones

El producto escalar es eficiente como medida de similaridad, cuando los patrones del conjunto de entrada tienen igual longitud, o parecida.

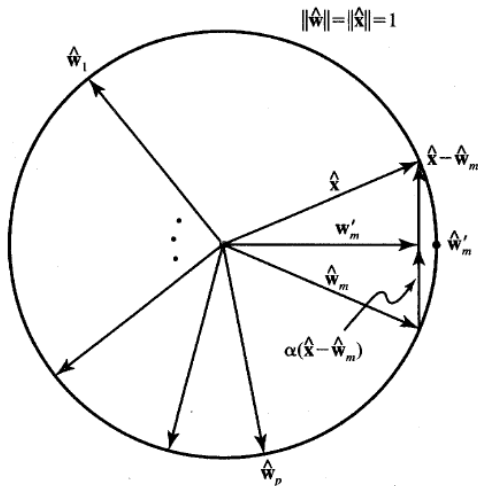
La distancia euclídea es una medida más general que la del producto escalar. Si los vectores están normalizados, la distancia euclídea es proporcional al producto escalar.

Aprendizaje Competitivo: Ejemplo



Red competitiva con p unidades de salida.

Regla de Aprendizaje Competitivo



Como $\hat{\mathbf{w}}_m$ es la neurona más similar a $\hat{\mathbf{x}}$ es la ganadora. El vector de pesos de $\hat{\mathbf{w}}_m$ se actualizará, para hacerse aún más similar al patrón de entrada $\hat{\mathbf{x}}$.

Aprendizaje competitivo: medida producto escalar

1. **Inicialización:** Los pesos de la red se inicializan en pequeños valores random. Normalizar los pesos de la red y los patrones de entrada.

Mientras haya cambios significativos en los pesos hacer:

- 1) **Elegir un patrón de entrada x al azar**
- 2) **Determinar la ganadora:** la neurona de la red que produce la respuesta máxima ante la entrada x es declarada ganadora. La ganadora entre todas las p neuronas que participan de la competencia es
$$w_m^t x = \max_{i=1,2,\dots,p}(w_i^t x)$$
- 3) **Actualización de los pesos:** La neurona ganadora m actualiza sus pesos según $\Delta w_m = \alpha(x - w_m)$, con $0 < \alpha < 1$
- 4) **Renormalización:** Normalizar los pesos de la ganadora.
- 5) **Decrementar el factor de aprendizaje:** $\alpha(t) = \alpha_0 t^{-\beta}$
(con $\beta \leq 1$)

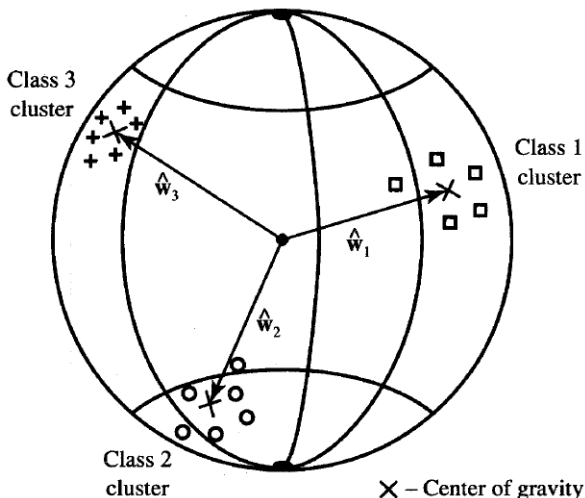
Aprendizaje competitivo: medida distancia euclidea

1. **Inicialización:** Se toman algunos patrones representativos del conjunto de entrenamiento y se inicializan los pesos de la red con ellos.

Mientras haya cambios significativos en los pesos hacer:

- 1) **Elegir un patrón de entrada x al azar**
- 2) **Determinar la ganadora:** La ganadora entre todas las p neuronas que participan de la competencia es
$$\min_{i=1,2,\dots,p} \|x - w_i\|$$
- 3) **Actualización de los pesos:** La neurona ganadora m actualiza sus pesos según $\Delta w_m = \alpha(x - w_m)$, con
 $0 < \alpha < 1$
- 4) **Decrementar el factor de aprendizaje:** $\alpha(t) = \alpha_0 t^{-\beta}$
(con $\beta \leq 1$)

Detección de Clusters: red con 3 unidades de salida



Tras el entrenamiento, los pesos de las unidades de salida de la red serán los centroides de los datos de entrada.

Aprendizaje Competitivo vs. Hebbiano: relación

Si hay una sola neurona ganadora, el output O_i de la red es 1 para la ganadora, y 0 para todas las demás. Entonces, podemos reescribir la regla de aprendizaje competitivo $\Delta w_m = \alpha(x - w_m)$ como:

$$\Delta w_{ij} = \alpha O_i (x_j - w_{ij})$$

que es idéntica a la regla de aprendizaje hebbiano generalizado de Oja:

$$\Delta w_{ij} = \alpha V_i \left(x_j - \sum_{k=1}^N V_k w_{kj} \right)$$

(Pero recordar que en aprendizaje hebbiano se actualizan todas las neuronas, y en competitivo la neurona ganadora)

Aprendizaje Competitivo: algunas variantes

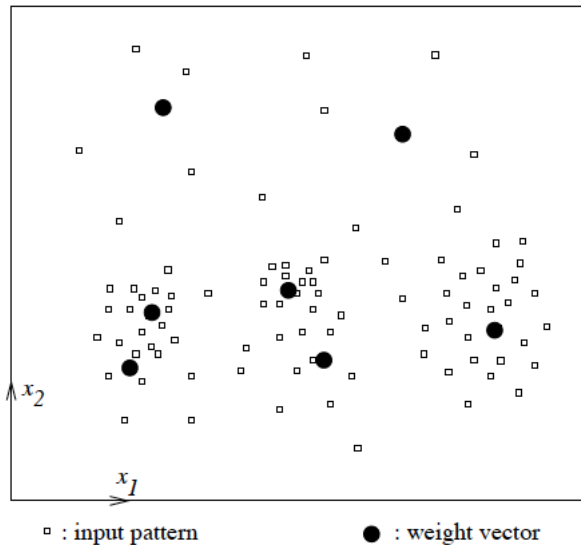
- ▷ *Variar la cantidad de neuronas de salida:* en lugar de un esquema fijo de neuronas de salida, eliminar las neuronas que no ganan casi nunca.
- ▷ *Función de vecindad excitatoria/inhibitoria:* actualizar la ganadora y las vecinas a ella, según cierta función de vecindad.
- ▷ *Variar la regla que determina la ganadora:* usar otras funciones de distancia al patrón de entrada, hacer que ganen varias neuronas en vez de una sola, etc.

- ☒ *Aprendizaje No Supervisado*
- ☒ *Aprendizaje Competitivo*
- ☐ *Aplicación: Vector Quantization*

Vector quantization

Una aplicación importante del aprendizaje competitivo es la cuantización vectorial. Un esquema de cuantización vectorial divide el espacio de entrada en un número de subespacios disjuntos y representa cada vector de entrada x con la etiqueta del subespacio en el cual cae (i.e., el índice k de la neurona ganadora). La diferencia con clustering es que no estamos tan interesados en encontrar clústeres de datos similares, sino en cuantizar el espacio de entrada. La densidad de las neuronas y en consecuencia de los subespacios es más alta en aquellas áreas en las cuáles los inputs tienen mayor densidad de probabilidad.

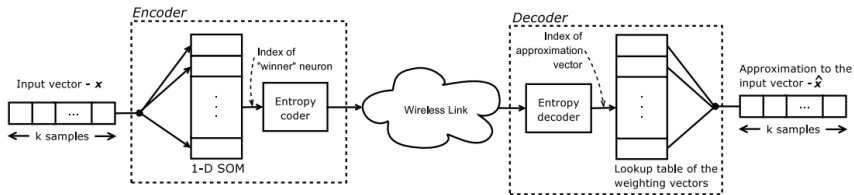
Vector quantization



Ejemplo de Vector quantization

Objetivo: compresión. El vector de entrada es reemplazado por el índice de la unidad de salida que dispara.

Los centroides son conocidos por el codificador y el decodificador. En lugar de transmitir toda la información, se transmite el índice del centroide más cercano. El decodificador usa el centroide como una aproximación del dato original.



Referencias

Fausett, *Fundamentals of neural networks*, 1994.

Haykin, *Neural Networks and Learning Machines*, 2009.

Hertz, *Introduction to the theory of neural computation*, 1991.

Krose y Smagt, *An Introduction to Neural Networks*, 1996.

LeCun, *The Next Frontier in AI: Unsupervised Learning*, 2016.

Zurada, *Introduction to artificial neural systems*, 1992.