

REDES NEURONALES ARTIFICIALES

Lección 5

REDES DE CREENCIA
PROFUNDA

Divergencia Contrastiva

El entrenamiento para las maquinas Restringidas de Boltzmann puede ser aplicado con el algoritmo de un solo paso de la divergencia contrastiva (**CD-1**), el cual funciona así:

1. FASE POSITIVA:

- Una muestra de entrada \mathbf{v} se sujeta a la capa visible.
- \mathbf{v} se propaga a la capa oculta de una manera similar como a las redes feedforward. El resultado de las activaciones de la capa oculta es \mathbf{h} .

2. FASE NEGATIVA:

- Propagar \mathbf{h} de regreso a la capa visible con resultado \mathbf{v}' (recordar que las conexiones entre las capas visibles y ocultas son bidireccionales).
- Propagar la nueva \mathbf{v}' de vuelta a la capa oculta con resultado de activaciones \mathbf{h}' .

Divergencia Contrastiva

3. Actualización de peso:

$$w(t+1) = w(t) + a (\mathbf{v}\mathbf{h}^\top - \mathbf{v}'\mathbf{h}'^\top)$$

Donde a es la tasa de aprendizaje y \mathbf{v} , \mathbf{v}' , \mathbf{h} , \mathbf{h}' , y \mathbf{w} son vectores.

- La fase positiva (\mathbf{h} dada \mathbf{v}) refleja la representación interna de la red de datos de la *vida real*.
- La fase negativa representa un intento de recrear los datos, basados en esta representación interna (\mathbf{v}' dada \mathbf{h}).

Divergencia Contrastiva

El objetivo principal es que los datos generados sean lo más cercano posible a la de la vida real y esto se refleja en la fórmula de actualización del peso.

En otras palabras, la red tiene una percepción de cómo los datos de entrada pueden ser representados, por lo que trata de reproducir los datos basándose en esta percepción. Si su reproducción no es lo suficientemente cercana a la realidad, hace un ajuste y lo intenta de nuevo.