

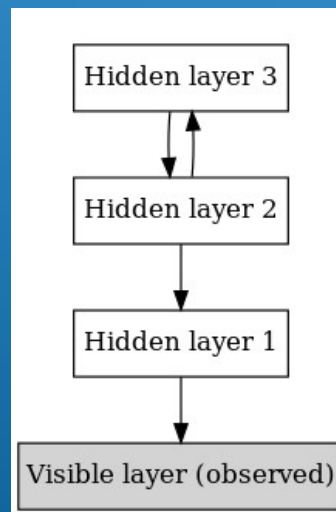
REDES NEURONALES ARTIFICIALES

Lección 5

REDES DE CREENCIA PROFUNDA

Definición

Ahora si podemos definir las redes de creencia profunda o deep belief networks (DBN), las cuales son simplemente la conglomeración de varias Máquinas de Boltzmann Restringidas, siendo así que cada capa oculta de una subred es la capa visible de la siguiente. El entrenamiento puede ser realizado individualmente para cada MBR que compone la red utilizando el algoritmo de divergencia contrastiva, conocido como greedy layer-wise, lo cual ayuda a superar el problema de sobre-entrenamiento y la desaparición de gradiente.



Representación grafica de una DBN

Entrenamiento

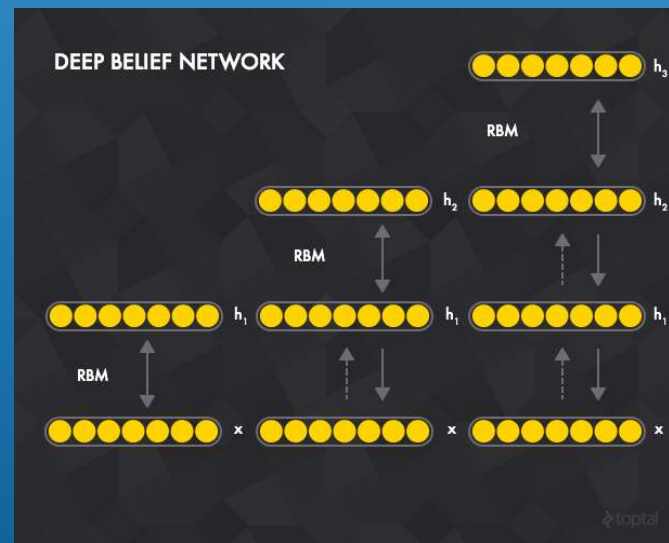
Como ya fue mencionado se entrenará una MRB a la vez hasta llegar al final de la red.

En este caso, la capa oculta de la MRB t actúa como una capa visible de MRB $t + 1$. La capa de entrada de la primera MRB es la capa de entrada para toda la red, y la greedy layer-wise de pre-entrenamiento funciona así:

1. Capacita la primera MRB $t=1$, utilizando divergencia contrastiva con todas las muestras de entrenamiento.
2. Capacita la segunda MRB $t=2$. Puesto que la capa visible para $t=2$ es la capa oculta de $t=1$, la formación comienza por sujetar la muestra de entrada a la capa visible de $t=1$, la cual se propaga hacia adelante a la capa oculta de $t=1$. Estos datos sirven, posteriormente, para iniciar el entrenamiento de divergencia contrastiva para $t=2$.

Entrenamiento

3. Repite el procedimiento anterior para todas las capas.
4. Después de pre-entrenar, la red se puede ampliar mediante la conexión de una o más capas completamente conectadas a la capa oculta final MBR. Esto forma un perceptrón de múltiples capas que puede ser afinado mediante propagación hacia atrás.



Representación grafica del entrenamiento de una DBN