

Mecanismo de Aprendizaje de una Red Neuronal (RN)

El mecanismo de aprendizaje de una RN se basa en la configuración simultánea de todos los parámetros internos a partir del procesamiento de un gran conjunto de datos, con su resultado correctamente etiquetado. Este mecanismo crea conexiones entre neuronas, destruye conexiones entre neuronas o modifica las conexiones preexistentes para optimizar los resultados y que sean más correctos en cada ciclo de prueba con cada dato disponible. Este mecanismo se llama retropropagación (backpropagation). Y funciona gracias al algoritmo matemático del descenso del gradiente que minimiza la función error en la salida.

Partes del mecanismo de aprendizaje de RN

Se requiere la RN, que es un conjunto de neuronas agrupadas en capas, los datos de entrenamiento, y los resultados. Los datos deben estar vinculados a sus resultados correctos, las capas de neuronas pueden ser de tres tipos: la capa de entrada, capas ocultas y la capa de salida. Cada neurona en una capa está conectada a todas las neuronas de la capa siguiente y anterior mediante conexiones ponderadas por pesos.

Capa de entrada: son las neuronas que se le inyecta directamente los valores de entrada de la red, y hacen el primer filtrado de datos, esta capa tiene obligatoriamente la misma cantidad de neuronas que la cantidad de datos se quieran analizar.

Capa Oculta: estas capas no están restringidas a un número de neuronas, cada neurona está conectada con todas las neuronas de la capa anterior, y la salida de cada neurona está conectada a todas las neuronas de la capa siguiente.

Capa de salida: esta capa tiene la misma cantidad de neuronas que cantidad de salidas, y cada neurona está conectada a todas las neuronas de la capa anterior, y la salida a una única predicción. Esta capa es la primera en ajustarse, porque el error se calcula de forma trivial, verificando si coincide la predicción con el resultado del entrenamiento.

Dato de entrada: estos datos son conjuntos de valores donde la red va a ser ajustada de manera precisa. Estos datos se conectan directamente a la capa de entrada en cada iteración de ajuste.

Dato de salida: para cada dato de entrada, se le suministra a la red su valor de salida, estos datos de salida se usan para calcular el error y corregir en la retropropagación a cada neurona.

Funcionamiento de una Neurona:

Cada neurona realiza dos operaciones fundamentales: la suma ponderada de las entradas y la aplicación de una función de activación. La suma ponderada se calcula multiplicando cada entrada (S_n) por su peso correspondiente (α_n) y sumando los resultados. La función de activación (típicamente la función sigmoide, que tiene bajo costo computacional) introduce no linealidades en el modelo, permitiendo que la red aprenda patrones más complejos.

La salida de cada neurona respeta la ecuación $S_t = \text{sigmoid}(\sum_{n=0}^{N, si n \neq t} \alpha_n \times S_n)$ y en mecanismo de backpropagation se ajustan los α_n de a un paso a la vez.

El mecanismo aprendizaje

Los α_n son ajustados cíclicamente de manera que la red aproxime la ecuación de la solución intrínseca en los datos de entrenamiento lo mejor posible, ya que la ecuación que representa la RN está codificada en los α_n que se ajustan para que la red resuelva con el menor error posible los datos de entrenamiento, y por interpolación resuelve para cualquier dato de entrada. En la Figura 1 finalmente se plasma este mecanismo de aprendizaje, donde los datos de entrada se procesan desde la capa de entrada hacia la capa de salida, los resultados de la capa de salida se comparan con los resultados correctos, y se calcula la corrección del error para cada peso " α_n " de cada capa.

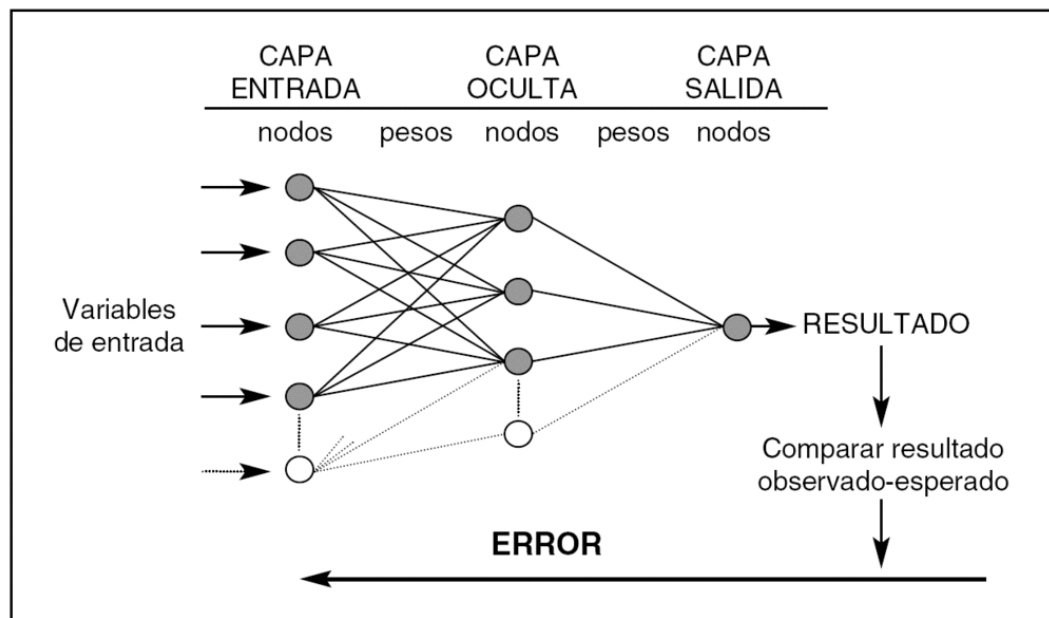


Figura 1: retropropagación ajuste del error en una RN