

Optimización en el entrenamiento del perceptrón multicapa

Expositor: Hevert Vivas

Resumen

Las *Redes Neuronales Artificiales* son sistemas de procesamiento de información que funcionan de manera similar a las redes neuronales biológicas. Estas redes tienen en común con el cerebro humano la distribución de las operaciones a realizar en una serie de elementos básicos que, por analogía con los sistemas biológicos, se denominan *neuronas artificiales* las cuales, están relacionadas entre sí, mediante una serie de conexiones llamadas *pesos sinápticos*. Estos pesos y conexiones varían con el tiempo mediante un proceso, usualmente iterativo, conocido como *aprendizaje o entrenamiento* de la red, en el cual, la determinación de los pesos se hace de tal manera que al final del proceso, la información *relevante* de los datos de entrada quede incorporada en la estructura de la red.

En el *entrenamiento*, está inmersa una *función error* que depende explícitamente de los pesos *sinápticos* y proporciona el error que comete la red, comparando en cada iteración, la salida obtenida con la salida deseada. Matemáticamente, la *función error* es un campo escalar y el *entrenamiento de la red* consiste en encontrar un vector (es decir una configuración de pesos) que *minimice* dicha *función error*. Por ello, asociado al rendimiento de una red neuronal, tenemos un *problema de minimización*, propiamente, un *problema de mínimos cuadrados no lineales*.

Dentro del grupo de modelos de redes neuronales artificiales está *el perceptrón multicapa*: una red neuronal *unidireccional* constituida por tres capas o más, cuyo aprendizaje se hace mediante un algoritmo denominado *retropropagación de errores (backpropagation)*. Esta red, junto con su algoritmo de aprendizaje, es el modelo más empleado en aplicaciones prácticas.

En nuestra investigación, proponemos e implementamos por primera vez, el *método secante estructurado* para el entrenamiento del *perceptrón multicapa* y analizamos su desempeño numérico comparándolo con métodos ampliamente usados con el mismo propósito. Pruebas numéricas preliminares muestran un buen desempeño numérico del método propuesto.