

Redes Neuronales

Gonzalo V. Castiglione, Alan E. Karpovsky, Martín Sturla
Estudiantes Instituto Tecnológico de Buenos Aires (ITBA)

12 de Abril de 2012

Entrega Preliminar 2 - Informe

Resumen—El presente informe busca analizar y comparar distintas implementaciones y arquitecturas de redes neuronales con aprendizaje supervisado que resuelvan el problema del *AND* y *OR* lógico para N bits con $2 \leq N \leq 5$ a través del uso de tres variantes de funciones de transferencia.

Palabras clave—perceptrón, función de transferencia, red neuronal, aprendizaje supervisado, conjunto de entrenamiento, conjunto de testeo

I. INTRODUCCIÓN

LAS redes neuronales son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales y los seres humanos. Se trata de un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida. Cabe destacarse que las redes neuronales no dejan de ser un modelo y es debido a esto que no tienen en cuenta muchas de las funciones y cualidades del cerebro humano.

Una red neuronal se compone de unidades llamadas neuronas. Cada neurona recibe una serie de entradas a través de interconexiones y emite una salida. Esta salida viene dada por dos funciones:

- **Función de excitación:** Por lo general consiste en la sumatoria de cada entrada multiplicada por el peso de su interconexión (*suma pesada de las entradas*). Si el peso de una conexión es positivo, la misma se denomina excitatoria; si es negativo, se denomina inhibitoria.
- **Función de transferencia:** Se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que queramos darle a dichas salidas. Algunas de las más utilizadas son la *función sigmoidea* (para obtener valores en el intervalo $[0, 1]$) y la *tangente hiperbólica* (para obtener valores en el intervalo $[-1, 1]$).

II. DESARROLLO

A. Modelado del problema

Se representó la red neuronal como una **matriz de pesos**. Cada neurona es una fila de pesos, cada capa de neuronas es una matriz de pesos, la red neuronal, por con-

siguiente, es un vector de matrices. Lo interesante es que hallar el valor de la red neuronal con un cierto input se reduce a multiplicar el vector input por cada una de estas matrices. A su vez esta implementación facilita el manejo de errores.

B. Features

Se implementaron tres tipos de funciones para modificar la variable η (*learn rate*): La primera de ellas es a valor **constante**, es decir que η nunca es modificada, la segunda es **annealed** que reduce η exponencialmente y por último se tiene un *learning rate* **dinámico** que lo que hace es ver si el error se reduce consistentemente y, en tal caso, incrementa el η sin superar un valor máximo preestablecido (0,5), caso contrario, η es reducida exponencialmente.

III. RESULTADOS

A. AND lógico

Se decidió, para la implementación del *AND*

Capas de la red	Función de activación	Épocas	Error de aprendizaje
[2;3;1]	Escalón	a	d
[2;3;1]	Lineal	b	e
[2;3;1]	Sigmoidea	c	f

TABLE I
COMPARACIÓN DE RED PARA AND LÓGICO

B. OR lógico

Se decidió, para la implementación del *AND*

Capas de la red	Función de activación	Épocas	Error de aprendizaje
[2;3;1]	Escalón	a	d
[2;3;1]	Lineal	b	e
[2;3;1]	Sigmoidea	c	f

TABLE II
COMPARACIÓN DE RED PARA OR LÓGICO

IV. CONCLUSIÓN

Es notable destacar