

# Chapter 2: Backpropagation

## Resumen del capitulo

En es capitulo analizamos a detalle las redes neuronales feedforward. Feedforward es un tipo de red neuronal, que por lo general es utilizada en el reconocimiento de patrones, su arquitectura radica en una capa de salida, una o mas capas ocultas y una capa de salida. Las conexiones entre capas van de un solo sentido, desde la capa de entrada hacia la capa de salida. En este tipo de redes es muy común la utilización del algoritmo de backpropagation para su entrenamiento, aun que no es obligatorio que sea de esa manera, se puede utilizar otro algoritmo de entrenamiento.

Dado que este tema ya lo hemos abordado solo puntualizo lo nuevo abordado en este capítulo.

1. Las redes feedforward usualmente tienen solo una capa oculta. Pocas veces o raras veces vemos redes con más de una capa.
2. Las funciones de activación compatibles con el algoritmo de entrenamiento backpropagation son: función sigmoide y tangente hiperbólica.
3. El numero de neuronas en la capa oculta es un número muy importante, y aunque no se aborda el tema de como calcular este valor se hace la observación que un número erróneo de capas ocultas podría provocar problemas (overfitting y underfitting).

Estas observaciones no habían sido mencionadas al momento, sin dejar de lado que el resto que conlleva las redes feedforward.

Además de explicar la parte teórica el autor no brinda una seria de clases que se implementan para la resolución del problema de una red neuronal feedforward.

## Preguntas de revisión

**1. ¿Qué es una función de activación? Explique cuándo podría utilizar la función de activación de tangente hiperbólica sobre la función de activación sigmoidea.**

Una función de activación representa la salida de una neurona especifica. Se calcula mediante la aplicación de cierta función (sigmoide, hiperbólica) al producto punto de los pesos y las entradas de dicha neurona.

Las funciones de activación sigmoide e hiperbólica se usan en casos muy distintos: la función sigmoide es de uso específico para las redes que solo usan valores positivos, y la hiperbólica se utiliza cuando son requeridos valores negativos.

**2. ¿Cómo se puede determinar si una función de activación es compatible con el método de entrenamiento Backpropagation?**

Para verificar si una función de activación es compatible con el algoritmo de entrenamiento backpropagation basta con calcular la derivada de la función, si el valor es distinto de un valor constante, es decir la derivada es otra función, la función de activación es compatible con el algoritmo de entrenamiento backpropagation.

**Ejemplo**



La función sigmoide es compatible

****

La función presentada no es compatible

**3. Considere una red neuronal con una neurona de salida y tres neuronas de entrada. Los pesos entre las tres neuronas de entrada y la neurona de salida son 0,1, 0,2 y 0,3. El umbral es 0,5. ¿Cuál sería el valor de salida para una entrada de 0,4, 0,5 y 0,6?**

Red neuronal

Cálculo de la salida

**dotProduc =** (0.4 \* 0.1 + 0.5 \* 0.2 + 0.6 \* 0.3) = 0.32

**FuncSigmoide =** 1 / ( 1 + exp( - 0.32 ) ) = 0.579324252

**Salida** **=** 0.579324252

**4. Explique el papel de la tasa de aprendizaje en la propagación hacia atrás.**

La velocidad en que una red neuronal aprende. Dentro del algoritmo de backpropagation la tasa de aprendizaje representa la longitud del ajuste que se debe realizar, es decir, indica que tanto debe ajustarse el peso en la iteración, de tal manera que el ajuste no sea tan grande ni tampoco tan pequeño, para el caso de que se muy grande la red aprenderá muy rápido y generara una tasa de error no deseada o en el caso de que se muy pequeño el aprendizaje será mínimo y la llegada al error deseado será muy lento. Usualmente los valores utilizados para la tasa de error son 0.01 o 0.001.

**5. Explique el papel del *momentum* en la propagación hacia atrás.**

Recuerda el incremento aplicado a las variables en cada iteración y determina la siguiente actualización como una combinación lineal entre el gradiente y el incremento anterior. Es decir, aplica a los incrementos cierta "inercia" de forma que varíen más lentamente.

## Vocabulario

**Activation Function**

Funciones utilizadas en redes neuronales, como por ejemplo en redes feetforwar, y son utilizadas con la finalidad de escalar los rangos de salida de una neurona a rangos adecuados. Generalmente son usadas las funciones de activación: Tangente hiperbólica y sigmoide, la utilización de una o de otra depende de los rangos de salida que se desean trabajar, siendo sigmoide adecuada para valores positivos y tangente hiperbólica adecuada para valores positivos y negativos.

**Backpropagation**

Algoritmo de entrenamiento utilizado en redes neuronales, generalmente es más utilizado junto a redes feedfoward, pero no es la única red en la cual se puede aplicar. El algoritmo de backpropagation es un entrenamiento supervisado donde se proveen las salidas idóneas para cada entrada proporcionada. El nombre de este entrenamiento proviene de su manera de actuar, iniciando la actualización de pesos desde la capa de salida hasta la capa de entrada, es decir, en retroceso.

**Derivative**

La derivada de una función se define como, resultado de un límite y representa, geométricamente, la pendiente de la recta tangente a la gráfica de la función en un punto. Hablando dentro del ámbito de redes neuronales la derivada no cambia de conceto, es utilizada dentro del algoritmo de entrenamiento backpropagation en el calculo de los valores delta, los cuales, son necesarios calcular para poder realizar el ajuste de pesos.

**Feedforward**

Tipo de red neuronal, que generalmente es utilizada en el reconocimiento de patrones. Este tipo de red se caracteriza porque sus conexiones van en un solo sentido (desde la capa de entrada a la capa de salida). La red neuronal feedforward generalmente es utilizada con el algoritmo de entrenamiento backpropagation, aunque no necesariamente debe ser de esta manera, pero por lo general así es.

**Hyperbolic Tangent Activation Function**

Función de activación que utiliza la función tangente hiperbólica y genera valores dentro del rango [-1, 1]. Esta función de activación es utilizada en redes neuronales cuando se requiere que su valor de salida se encuentre dentro de rango ya mencionado.

**Sigmoid Activation Funtion**

Función de activación que utiliza la función sigmoide y genera valores dentro del rango [0,1]. Esta función de activación es utilizada en redes neuronales cuando se requiere que su valor de salida se encuentre dentro del rango ya mencionado.

**Linear Activation Function**

Función de activación que utiliza una función lineal, generalmente de primer grado, por lo que genera un amplio rango de valores. Esta función de activación es utilizada en redes neuronales cuando se requiere que su valor de salida sea presenta un rango de valores.

**Learning Rate**

Valor numérico que indica la velocidad con la que la red neuronal aprende. Cuando hablamos de la velocidad de aprendizaje nos referimos a la longitud que se debe realizar en el ajuste de pesos. Esta longitud no debe ser ni muy grande ni muy pequeño, para evitar este tipo de casos se utilizan valores como 0.01 y 0.001.

**Momentum**

Recuerda el incremento aplicado a las variables en cada iteración y determina la siguiente actualización como una combinación lineal entre el gradiente y el incremento anterior. Es decir, aplica a los incrementos cierta "inercia" de forma que varíen más lentamente.

**Overfitting**

Ocurre cuando en la capa oculta se cuenta con un numero elevado de neuronas se produce el sobreajuste. En este punto la red tiene mucha capacidad de procesamiento de información que la información de entrada la limita.

**Underfitting**

Ocurre cuando en la capa oculta no se cuentan con las neuronas suficientes para detectar de manera adecuada las señales de un conjunto de datos complicado.