## 1

Capítulo 1

Activación de la red neuronal

* Suma
* Cálculo de la activación
* Funciones de activación
* Neuronas de sesgo

En este capítulo verá cómo calcular la salida para una red neuronal feedforward. La mayoría de las redes neuronales se basan de alguna manera en la red neuronal feedforward. Ver cómo se calcula esta simple red neuronal será la base para entender el entrenamiento y otras características más complejas de las redes neuronales.

En este capítulo se introducirán varios términos matemáticos. Se le mostrará notación de sumatoria y notación de fórmula matemática simple. Comenzaremos con una revisión del operador de sumatoria.

# Comprensión del operador de suma

En esta sección, echaremos un vistazo rápido al operador de suma. El operador de suma, representado por la letra griega de capital sigma se puede ver en la ecuación 1.1.

10



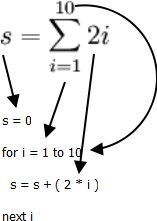
*s* = 2*i* (1.1)

*i*=1

## 2 Activación de la red neuronal

La ecuación anterior es un resumen. Si no está familiarizado con la notación sigma, es exactamente lo mismo que en programación en bucle **for**. La Figura 1.1 muestra la Ecuación 1.1 reducida al pseudocódigo.

**Figura 1.1:** Operador de suma al código



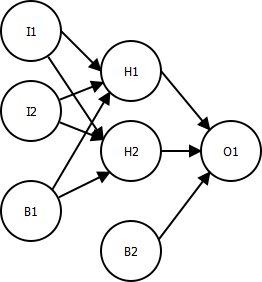
Como puede ver, el operador de suma es muy similar a un bucle **for.** La información justo debajo del símbolo sigma especie el valor indicado y la variable de indexación. La información por encima de la especie sigma el límite del bucle. La información a la derecha de sigma especifica el valor que se está sumando.

# Cálculo de una red neuronal

Comenzaremos analizando cómo una red neuronal calcula su salida. Ya debe conocer la estructura de una red neuronal a partir de la introducción del libro. Considere una red neuronal como la de la Figura 1.2.

## Cálculo de una red neuronal 3

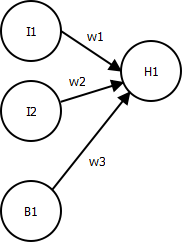
**Figura 1.2:** Una red neuronal simple



Esta red neuronal tiene una neurona de salida. Como resultado, tendrá un valor de salida. Para calcular el valor de esta neurona de salida (**O1**), debemos calcular la activación para cada una de las entradas en **O1**. Las entradas que se alimentan en **O1**  son **H1,** **H2**  y **B2.** La activación para **B2**  es simplemente 1.0, porque es una neurona de sesgo. Sin embargo, **H1**  y **H2** deben calcularse de forma independiente. Para calcular **H1**  y **H2,** se deben considerar las activaciones de **I1,** **I2**  y **B1.**  Aunque **H1**  y **H2** comparten las mismas entradas, no calcularán la misma activación. Esto se debe a que tienen diferentes pesos. Los pesos se represan por las líneas en el diagrama anterior.

Primero debemos ver cómo se hace un cálculo de activación. Este mismo cálculo de activación se puede aplicar a los demás cálculos de activación. Examinaremos cómo se calcula **H1.** La Figura 1.3 muestra sólo las entradas a  **H1**.

**Figura 1.3:** Cálculo de la activación de H1Šs



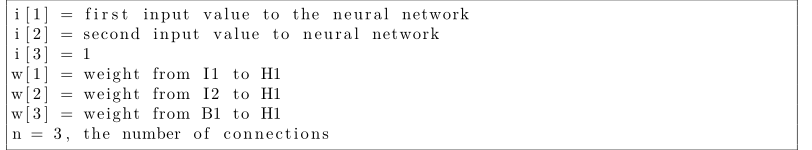
## 4 Activación de la red neuronal

Ahora veremos cómo calcular H1. Esta ecuación relativamente simple se muestra en la ecuación 1.2.



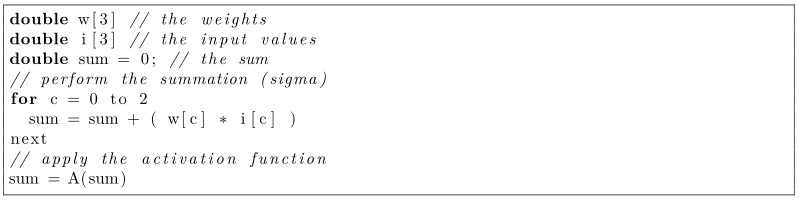
Para entender la Ecuación 1.2 primero examinamos las variables que entran en ella. Para la ecuación anterior tenemos tres valores de entrada, dados por la variable  **i**. Los tres valores de entrada son valores de entrada de  **I1,**  **I2** y  **B1**. **I1** e  **I2** son simplemente los valores de entrada que se proporcionó a la red neuronal para calcular la salida. **B1** es siempre 1, porque es la neurona de sesgo.

allí Son Además Tres Peso Valores Considera **w1**, **w2** Y **w3**. Estos Son el Ponderado conexiones entre **H1** y la capa anterior. Por lo tanto, las variables de esta ecuación Son:



Aunque la neurona de sesgo no es realmente parte de la matriz de entrada, uno siempre se coloca en la matriz de entrada para la neurona de sesgo. Tratar el sesgo como una neurona de solo avance hace que el cálculo sea mucho más fácil.

Para Entender Ecuación 1.2 Nosotros será Considerar eso Como Pseudocódigo.



Aquí resumemos cada una de las entradas veces su peso respectivo. Por último, esta suma se pasa a una función de activación. Las funciones de activación son un concepto muy importante en la programación de redes neuronales. En la siguiente sección examinaremos las funciones de activación.

* 1. **Funciones**de activación **5**

# Funciones de activación

Las funciones de activación se utilizan muy comúnmente en redes neuronales. Las funciones de activación cumplen varias funciones importantes para una red neuronal. La razón principal para utilizar una función de activación es introducir la no linealidad en la red neuronal. Sin esta no linealidad, una red neuronal podría hacer poco para aprender funciones no lineales. La salida que esperamos que las redes neuronales aprendan rara vez está cerca.

Las dos funciones de activación más comunes son la función de ac- tivación tangente sigmoide e hiperbólica. La función de activación tangente hiperbólica es la más común de estas dos, al igual que un rango de números de -1 a 1, en comparación con la función sigmoide which es sólo de 0 a 1.

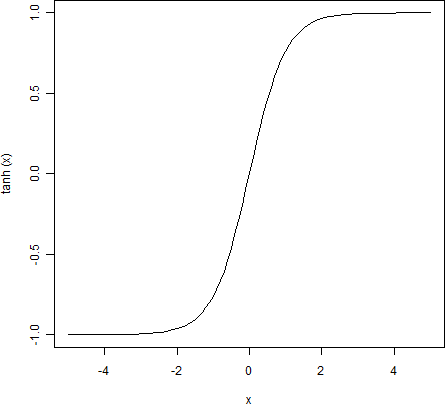
2*x*

*e* − 1

*f* (*x*) =  *e*2*x* + 1 (1.3)

La función tangente hiperbólica es en realidad una función trigonométrica. Sin embargo, nuestro uso para él no tiene nada que ver con la trigonomia. Esta función fue elegida para la forma de su gráfico. Puede ver un gráfico de la función tangente hiperbólica en la Figura 1.4.

**Figura 1.4:** La función tangente hiperbólica



¿Nota que el rango es de -1 a 1? Esto le da una gama mucho más amplia de números que puede

## 6 Activación de la red neuronal

Aceptar. También observe cómo los valores más allá de -1 a 1 se escalan rápidamente? Esto proporciona un rango constante de números para la red.

Ahora veremos la función sigmoide. Puedes ver esto en la Ecuación 1.4.

1

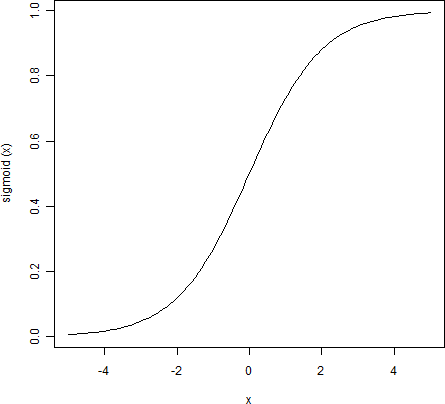
*f* (*x*) =

1 +  *e−x*

(1.4)

La función sigmoide también se denomina función logística. Normalmente no funciona tan bien como la función tangente hiperbólica. Sin embargo, si tiene todos los valores positivos en los datos de entrenamiento, puede funcionar bien. El gráfico de la función sigmoide se muestra en la Figura 1.5.

**Figura 1.5:** La función sigmoide



Como puede ver, escala los números a 1,0. También tiene un rango que sólo incluye números positivos. Es menos general que la tangente hiperbólica, pero puede ser útil. Por ejemplo, el tor de ópera XOR, como se describe en la introducción, sólo tiene números positivos. La función sigmoide supera la función tangente hiperbólica.

* 1. **Neuronas de sesgo 7**

# Neuronas de sesgo

¿Puede que te preguntes por qué se necesitan valores sesgados? Los valores de sesgo permiten que una red neuronal genere un valor de cero incluso cuando la entrada está cerca de una. La adición de un sesgo permite que la salida de la función de activación se desplace a la izquierda o a la derecha en el eje x. Para ver esto, considere una simple red neuronal donde una sola neurona de entrada **I1** está conectada directamente a una neurona de salida  **O1**. La red mostrada en el cuadro 1.6 no tiene sesgo.

**Figura 1.6:** Una conexión sin sesgo



La salida de esta red se calcula multiplicando la entrada (**x**) por el peso (**w**). A continuación, el resultado es pasar una función de activación. En este caso, estamos utilizando la función de activación sigmoide.

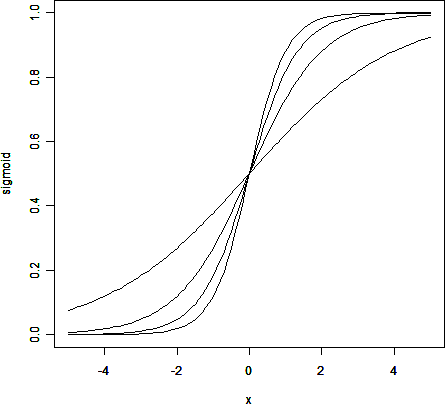
Considerar el Salida De el Sigmoideo Función Para el Siguientes Cuatro Pesos.



Dados los pesos anteriores, la salida del sigmoide será como se ve en la Figura 1.7.

## 8 Activación de la red neuronal

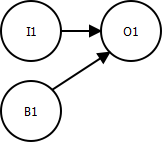
**Figura 1.7:** Ajuste de pesas



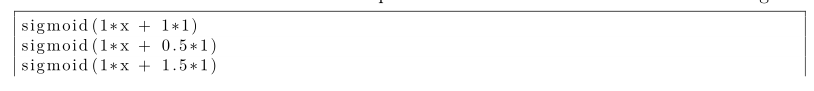
Cambiar el peso w altera la "inclinación"de la función sigmoide. Esto permite que la red neuronal aprenda patrones. Sin embargo, ¿qué pasa si desea que la red genere 0 cuando x es un valor distinto de 0, tal como 3? Sólo cambiar la inclinación del sigmoide wilno lograré esto. Debe ser capaz de desplazar toda la curva hacia la derecha.

Ese es el propósito del sesgo. La adición de una neurona de sesgo hace que la red neuronal aparezca como figura 1.8.

**Figura 1.8:** Una conexión sesgada



Ahora Nosotros Calcular Con el predisposición Neurona Presente. Nosotros será Calcular Para Varios predisposición Pesos.

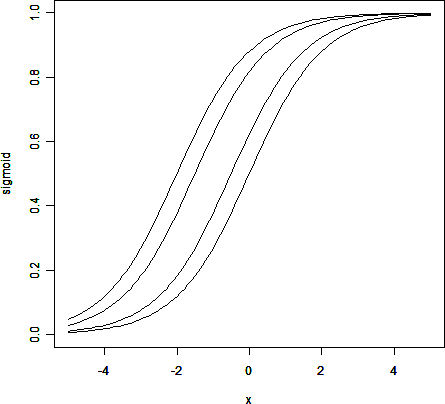


## Resumendel capítulo 9

## 

## éste Produce el Siguientes conspirar Visto En Figura 1.9.

**Figura 1.9:** Ajuste del sesgo



Como puede ver, toda la curva ahora cambia.

# 1.5 Resumen del capítulo

Este capítulo demostró cómo una red neuronal feedforward calcula la salida. La salida de una red neuronal se determina calculando cada capa sucesiva, después de la capa de entrada. La salida final de la red neuronal finalmente llega a la salida.

Las redes neuronales hacen uso de funciones de activación. Una función de activación proporciona no linealidad a la red neuronal. Dado que la mayoría de los datos que una red neuronal busca aprender no son lineales, las funciones de activación deben ser no lineales. Una activatien la función se aplica después de que se han multiplicado los pesos y las activaciones.

La mayoría de las redes neuronales tienen neuronas de sesgo. El sesgo es un concepto importante para las redes neuronales. Las neuronas de sesgo se agregan a todas las capas que no son de salida de la red neuronal. Las neuronas de sesgo son diferentes de las neuronas ordinarias de dos maneras muy importantes. En primer lugar, la salida de un

## 10 Activación de la red neuronal

la neurona sesgada siempre es una. En segundo lugar, una neurona sesgada no tiene conexiones entrantes. El valor constante de uno permite que la capa a responda con valores no nulos incluso cuando la entrada a la capa es cero. Esto puede ser muy importante para ciertos conjuntos de datos.

Las redes neuronales generarán valores determinados por los pesos de las conexiones. Estos pesos suelen establecerseen valores iniciales aleatorios. El entrenamiento es el proceso donde estos pesos aleatorios se ajustan para producir resultados significativos. Necesitamos una manera para que la red neuronal mida la eficacia de la red neuronal. Esta medida se denomina ca lculation de error. El cálculo de errores se describe en el siguiente capítulo.

## 1.5 Resumen del capítulo 11