



Inicio



Mi red



Empleos



Mensajes



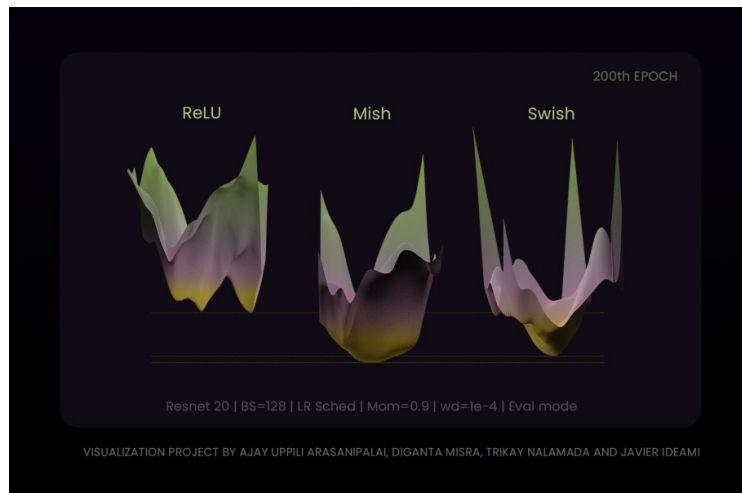
Notificaciones



Yo



Para negocios ▾ Publicar un anuncio de empleo gratis



La importancia de las funciones de activación en una red neuronal



Jorge Calvo Martin

Docente - Ingeniero Informático -
Postgrado Seguridad de la información y...

8 artículos

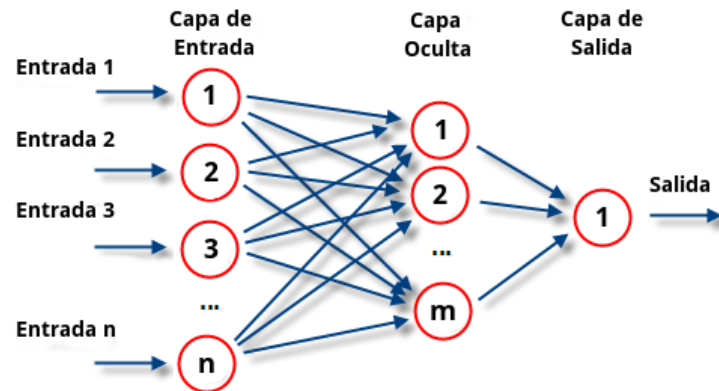
+ Seguir

18 de agosto de 2022

Para entender mejor este artículo y llegar a conocer como *unas funciones que están en el contenido de un curso de secundaria pueden llegar a ser tan importantes en conceptos tan complejos como el aprendizaje de una red neuronal.* Vamos empezar por entender que es una red neuronal artificial

Una red neuronal se define como un sistema que consiste en una serie de elementos todos interconectados, llamados "neuronas", que se organizan en capas que procesan la información utilizando respuestas de estado dinámico a entradas externas.

En el contexto de esta estructura, la capa de entrada introduce patrones en la red neuronal que tiene una neurona para cada componente presente en los datos de entrada y se comunica a una o más capas ocultas, se considera capas ocultas a todas las capas de la red exceptuando la de entrada y salida. Es en las capas ocultas donde sucede todo el procesamiento, a través de un sistema de conexiones caracterizado por pesos y sesgos (comúnmente conocidos como W y b)



¿Como funciona?

Con el valor de entrada que recibe la neurona se calcula una suma ponderada agregando también el sesgo y de acuerdo con el resultado y una **función de activación** preestablecida (las cuales veremos a continuación), se decide la activación o excitación de la neurona.

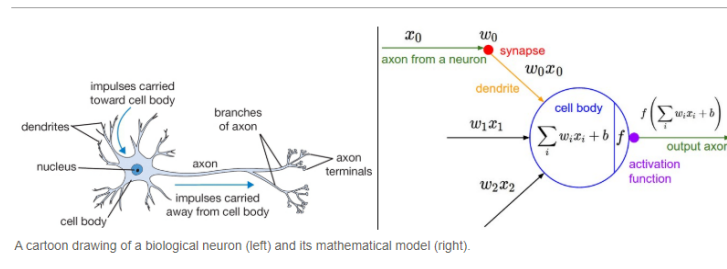
Posteriormente, la neurona transmite la información a otras neuronas conectadas en un proceso llamado "PassFoward". Al final de este proceso, la última capa oculta está vinculada a la capa de salida que tiene una neurona para cada posible salida deseada.

Una vez contextualizado el concepto de red neuronal,

¿Qué es realmente una función de activación?

Es una función matemática de la forma $f(x)$, que se agrega a una red neuronal artificial para ayudar a la red a aprender patrones complejos en los datos. Cuando se compara con un modelo basado en neuronas que está en nuestros cerebros, la función de activación al final decide qué se disparará a la siguiente neurona. Eso es

exactamente lo que hace una función de activación en una **ANN (Artificial neural network)** también. Toma la señal de salida de la celda anterior y la convierte en alguna forma que pueda tomarse como entrada para la siguiente celda. La comparación se puede resumir en la siguiente figura.



Fuente: cs231n by Stanford

¿Por qué son tan importantes?

Además de la similitud biológica que se habló anteriormente, también ayudan a mantener el valor de la salida de la neurona restringida a un cierto límite según nuestro requisito. Esto es importante porque la entrada a la función de activación es **$W \cdot x + b$** donde W son los pesos de la celda y x son las entradas para posteriormente sumar el sesgo b . Este valor, si no se restringe a un cierto límite, puede tener una magnitud muy alta, especialmente en el caso de redes neuronales muy profundas que tienen millones de parámetros. Esto dará lugar a problemas de cálculo. Por ejemplo, hay algunas funciones de activación como la sigmoide, la cual veremos a continuación, que generan valores específicos para diferentes valores de entrada entre 0 y 1.

La característica más importante de una función de activación es su capacidad para agregar no linealidad a una red neuronal.

¿Vale cualquier función en una red neuronal?

Realmente la respuesta es no, Las funciones de activación se pueden dividir básicamente en 2 tipos:

- **Función de activación lineal**: Las cuales ya no se usan en el Deep Learning, ya que la salida de las funciones no estará confinada entre ningún rango y la suma de diferentes funciones lineales sigue siendo una función lineal acotando así la activación de la neurona.
- **Funciones de activación no lineal**: Son las usadas en las redes neuronales, como veremos a continuación estas funciones permiten un acotamiento de los datos de salida. Algunos ejemplos son la función sigmoide o tangente hiperbólica

¿Cuales son las funciones de activación más usadas?

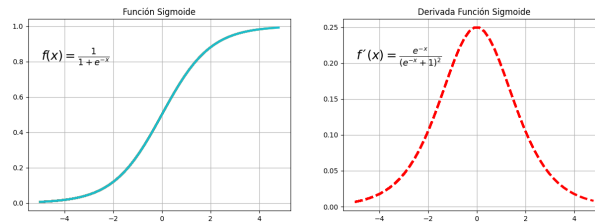
Hay varios tipos de funciones de activación usadas en el aprendizaje profundo y dependiendo de cual es el problema a resolver o el tipo de red usada, se usará una u otra función de activación. Incluso en muchas redes neuronales se usan varios tipos de funciones de activación a la vez (en diferentes capas).

Tipos de funciones de activación (las más usadas)

En todos los tipos de funciones mostradas tambien dibujaremos sus derivadas, ya que es un aspecto fundamental cuando una red neuronal realiza el proceso de Backpropagation (puedes saber más sobre este concepto en mi post: "[Backpropagation y gradiente descendente](#)")

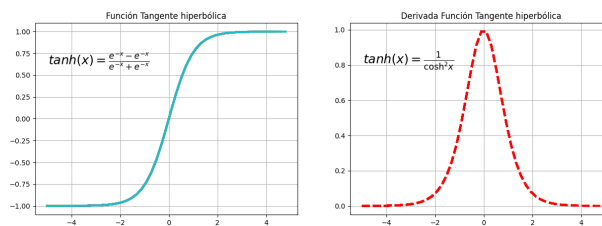
Función Sigmoide

La razón principal por la que usamos la función sigmoide es porque existe entre (0 a 1). Por lo tanto, se usa especialmente para modelos en los que tenemos que predecir la probabilidad como un resultado. Dado que la probabilidad de cualquier cosa existe solo entre el rango de 0 y 1.



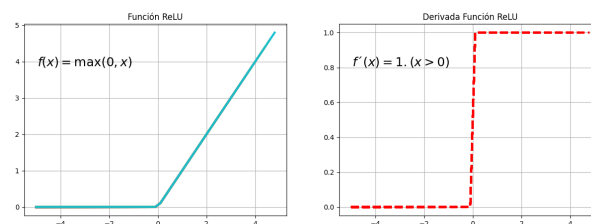
Función Tangente hiperbólica o Gaussiana

Es una función similar a la Sigmoide pero produce salidas en escala de $[-1, +1]$. Además, es una función continua. En otras palabras, la función produce resultados para cada valor de x .



Función RELU (Rectified Lineal Unit)

- ReLU es la función de activación más utilizada en el mundo en este momento. Desde entonces, se utiliza en casi todas las redes neuronales convolucionales o el aprendizaje profundo.
- Como puedes ver, ReLU está medio rectificado (desde abajo). $f(z)$ es cero cuando z es menor que cero y $f(z)$ es igual a z cuando z es superior o igual a cero.
- Es una función usada en las capas ocultas de nuestra red neuronal, NO en las de salida

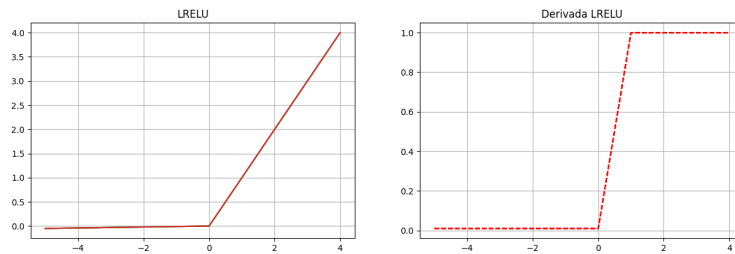


Leaky-ReLU

Leaky-ReLU es una mejora del valor predeterminado principal de ReLU, en el sentido de que puede manejar los

valores negativos bastante bien, pero aún presenta no linealidad.

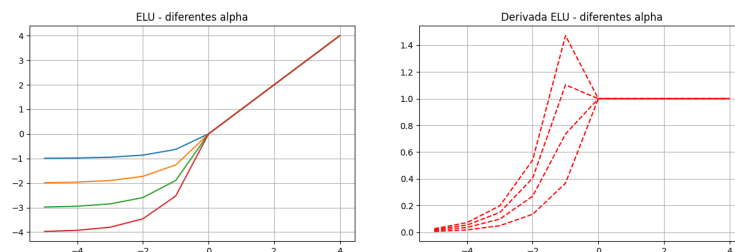
$$LRelu(x) = \max(0.01x, x)$$



Función ELU (Exponential Linear Unit)

La Unidad Lineal Exponencial (ELU) es una función de activación para redes neuronales. A diferencia de las ReLU, las ELU tienen valores negativos, lo que les permite acercar las activaciones de unidades medias a cero, como la normalización por lotes, pero con una menor complejidad computacional. Los cambios medios hacia cero aceleran el aprendizaje al acercar el gradiente normal al gradiente natural de la unidad debido a un efecto de cambio de sesgo reducido. Si bien las LReLU y las PReLU también tienen valores negativos, no garantizan un estado de desactivación resistente al ruido. Las ELU se saturan a un valor negativo con entradas más pequeñas y, por lo tanto, disminuyen la variación y la información propagadas hacia adelante.

Ref: <https://arxiv.org/pdf/1511.07289v5.pdf>

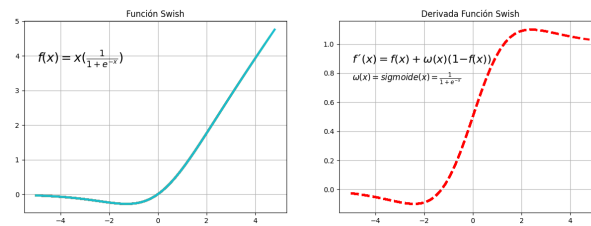


Función Swish

Swish es una función suave y no monótona que iguala o supera constantemente a ReLU en redes profundas aplicadas a una variedad de dominios desafiantes, como la clasificación de imágenes y la traducción automática. Es

ilimitado arriba y acotado abajo y es el atributo no monótono el que realmente crea la diferencia.

Ref: [https://arxiv.org/pdf/1710.05941v1.pdf?](https://arxiv.org/pdf/1710.05941v1.pdf?source=post_page)

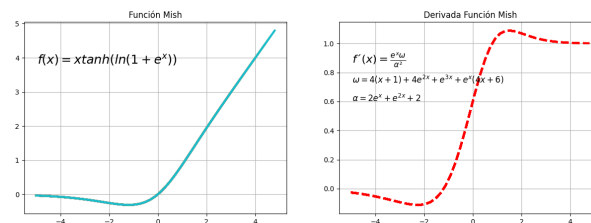


Función MISH

Mish, una nueva función de activación no monotónica autorregulada inspirada en la propiedad de activación automática de Swish.

Mish tiende a igualar o mejorar el rendimiento de las arquitecturas de redes neuronales en comparación con Swish, ReLU y Leaky ReLU a través de diferentes tareas en Computer Vision.

ReF: <https://arxiv.org/pdf/1908.08681.pdf>



Como conclusión se puede entender que debajo del aprendizaje de estos complejos modelos no hay más que matemáticas que debe ser entendida, o por lo menos conocida para aplicar los conceptos de la forma más productiva.

Como referencia os dejo este artículo donde se explica el funcionamiento de una red neuronal completa desde el punto de vista matemático.

<https://www.europeanvalley.es/noticias/crear-red-neuronal-desde-las-matematicas/>

Publicado por



Jorge Calvo Martin

Docente - Ingeniero Informático - Postgrado Seguridad de la información y C...
Fecha de publicación: 10 meses

8

artículos

+ Seguir

Me apasiona entender las similitudes que pueden existir entre lo humano y lo artificial, entre lo biológico y las matemáticas. En este caso en particular poder comprender como una simple función matemática es la encargada de activar las neuronas artificiales para que complejas redes puedan entrenar y aprender. Al igual que nuestras neuronas se activan por medio de sustancias químicas que las hacen transmitir pequeñas corrientes eléctricas que son los impulsos nerviosos, las funciones de activación son capaces de acotar o asignar probabilidades a los datos que entran en esas neuronas artificiales para transmitirlos a otras. Es cierto que hay conceptos mucho más complejos, el entender los modelos de aprendizaje profundo más allá de "una caja negra" (meto una imagen y me dice si es un perro o un gato, sin saber muy bien lo que ha realizado dentro para aprenderlo), es muy importante. Si esa caja negra, la podemos convertir en un poco más gris podremos ser más productivos y eficientes con esta tecnología. [#Deeplearning](#) [#machinelearning](#) [#matemáticas](#) [#ia](#) [#educacióndigital](#)



Recomendar



Comentar



Compartir



41

Reacciones



+29

0 comentarios



Añadir un comentario...



Jorge Calvo Martin

Docente - Ingeniero Informático - Postgrado Seguridad de la información y Ciberseguridad - Master AI & Datascience Education.

+ Seguir

Más de Jorge Calvo Martin



El Aumento sintético de datos.

Jorge Calvo Martin en Linke...



La debilidad matemática del algoritmo más seguro (RSA)

Jorge Calvo Martin en Linke...





Aprendizaje por Refuerzo en una IA




Jorge Calvo Martin en Linke...

[Ver todos los artículos \(8\)](#)

- Acerca de
- Pautas comunitarias
- Privacidad y condiciones ▾
- Sales Solutions
- Centro de seguridad

- Accesibilidad
- Empleo
- Opciones de publicidad
- Móvil

- Talent Solutions
- Marketing Solutions
- Publicidad
- Small Business

-  **¿Tienes preguntas?**
Visita nuestro Centro de ayuda.
-  **Gestiona tu cuenta y la privacidad**
Ve a los ajustes.
-  **Transparencia de las recomendaciones**
Más información sobre el contenido recomendado.

Seleccionar idioma

Español (Spanish)