

NOMBRE: Carlos Obed Cab Tun

MATRICULA: 7013

Horizontal

4 Es un criterio para detener el proceso de entrenamiento de una red neuronal.

6 Este optimizador por su simplicidad lo convierte en una buena opción para redes poco profundas.

8 Este optimizador proporciona tanto una tasa de aprendizaje adaptiva como la estimación adaptativa de momentos de primer y segundo orden.

10 Son algoritmos o métodos que se utilizan para cambiar los atributos de su red neuronal, como los pesos y la tasa de aprendizaje con el objetivo de reducir la pérdida (error).

Vertical

1 Es el proceso de aprendizaje en una red neuronal (Mayor costo computacional).

2 Luego de entrenar y validar la red neural, ésta se implementa para que ejecute la inferencia.

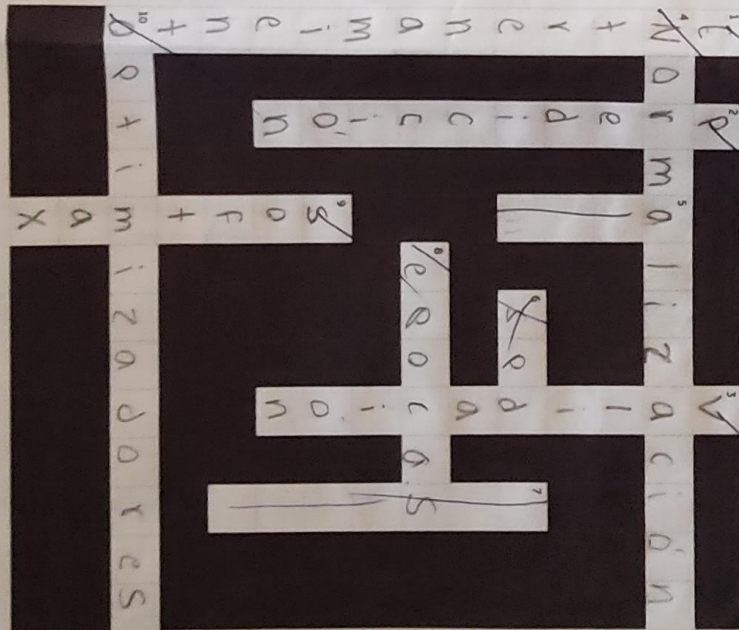
3 Es el proceso de comprobar si la red neuronal puede resolver nuevos problemas para los que ha sido entrenada.

5 Es una etapa para el preprocesamiento de datos se utiliza StandardScaler de Scikit-learn.

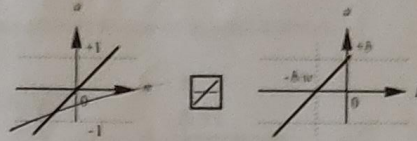
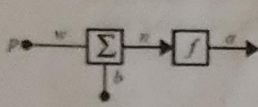
7 Para este optimizador la tasa de aprendizaje es adaptativo y es un optimizador muy robusto.

9 Esta función de activación transforma las salidas a una representación en forma de probabilidades, de tal manera que el sumatorio de todas las probabilidades de las salidas de 1.

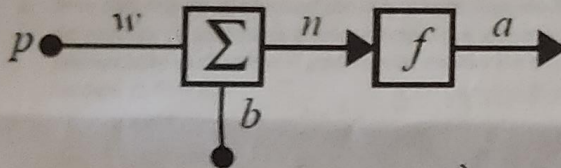
CRUCIGRAMA: 5 PUNTOS



1. Es muy parecido a un peso, excepto a que está asociada a una entrada constante 1. El modelo se volverá más flexible. (1 PUNTO) *Bias*

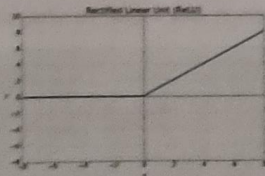


2. Escribe la ecuación de la siguiente neurona. (1 PUNTO)

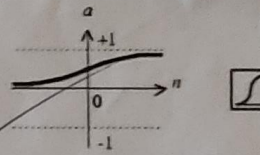


$$f(w \cdot p + b) = a$$

3. Se utilizan para introducir la no linealidad en las redes neuronales. (1 PUNTO)



Función de Activación o transferencia



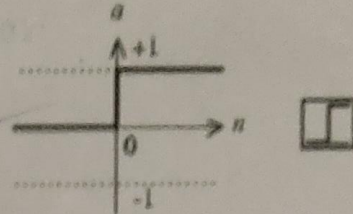
4. Esta función de activación tiene las siguientes características. (1 PUNTO)

- Establece la salida de la neurona en 0 si el argumento de la función es menor que 0, o 1 si su argumento es mayor o igual que 0.
- Esta función se utiliza para clasificar las entradas en dos categorías distintas.
- La salida de la neurona es discreta.

Función de Activación Escalon

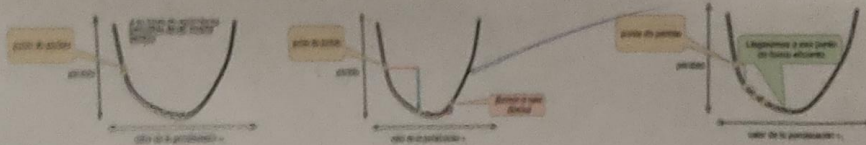
$$a = 0 \quad n < 0$$

$$a = 1 \quad n \geq 0$$



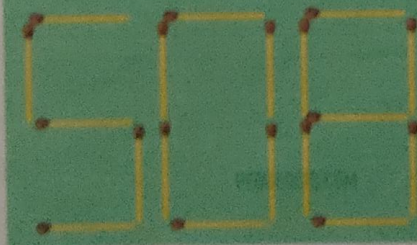
5. sirve para regular la velocidad con la que se realizan modificaciones incrementales de los pesos del modelo. En general, se emplea siempre que utilizemos una función de error o pérdida diferenciable y utilizemos el gradiente de esa función de error para corregir los pesos del modelo. (1 PUNTO)

Tasa de aprendizaje



PREGUNTA DE SALVACIÓN: (1 PUNTO)

¿CUAL ES EL MAYOR NÚMERO QUE PUEDE OBTENERSE MOVIENDO SOLO DOS CERILLAS?



990