Inicialización de parámetros

Las estrategias de inicialización de parámetros en redes neuronales son métodos para asignar valores iniciales a los pesos de la red antes de comenzar el proceso de entrenamiento. La elección adecuada de la inicialización puede tener un impacto significativo en el rendimiento y la convergencia de la red.

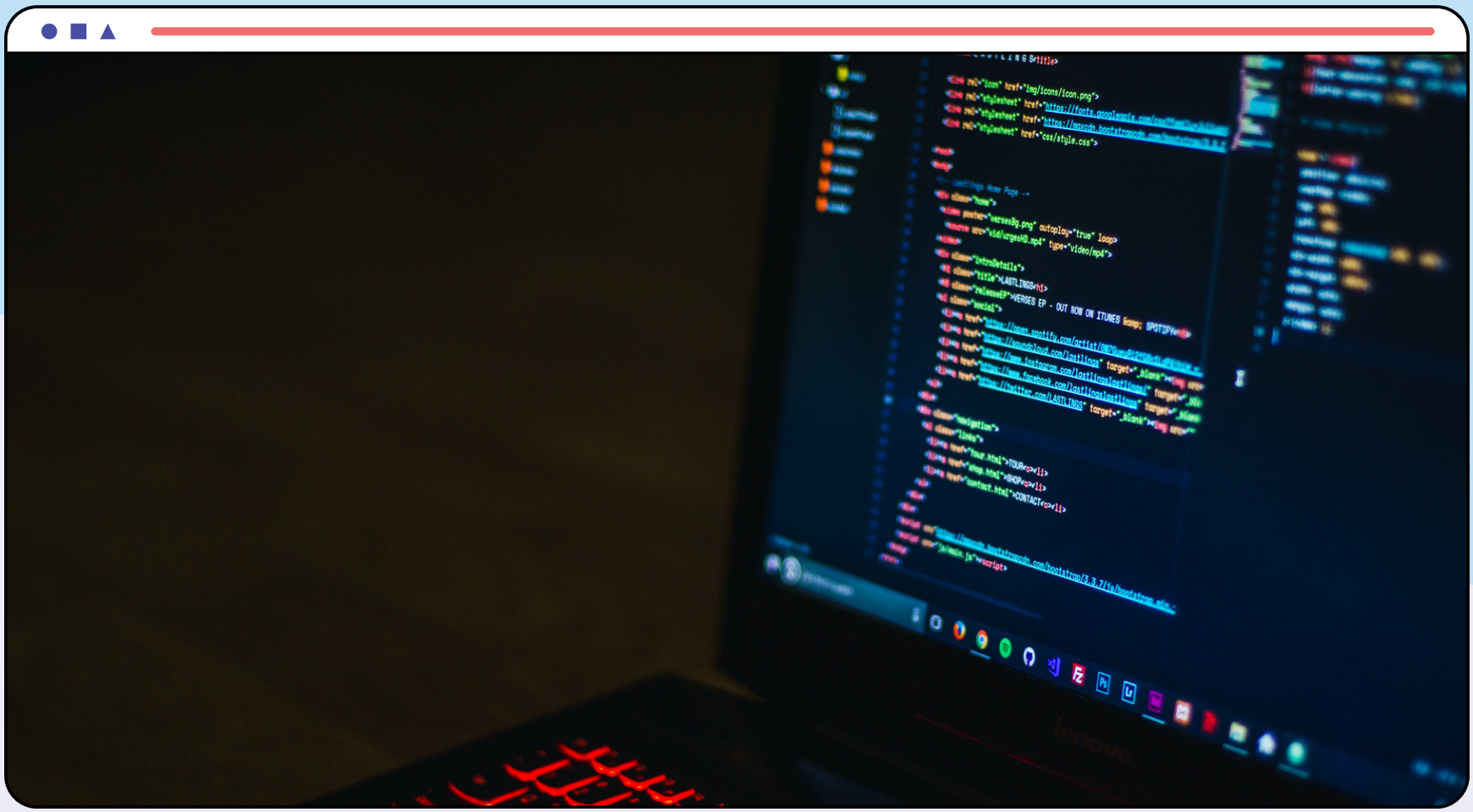
Una estrategia común es:

Inicialización aleatoria

Asigna pesos aleatorios a cada neurona en la red. Puede ser útil para romper simetrías, evitando que todas las neuronas aprendan lo mismo.

Por ejemplo: inicialización uniforme o gaussiana.

Las estrategias buscan establecer un equilibrio entre evitar gradientes demasiado pequeños o grandes al comienzo del entrenamiento, lo que puede conducir a problemas como el desvanecimiento o la explosión del gradiente.



Importancia de la Inicialización

- Evitar gradientes pequeños al inicio que pueden llevar a convergencia lenta o estancamiento.
- Evitar gradientes grandes que pueden causar problemas de explosión del gradiente.
- Promover una convergencia más rápida y estable durante el entrenamiento.
- La elección de la estrategia de inicialización dependerá del tipo de activaciones utilizadas en la red y puede requerir ajustes específicos según la arquitectura y la tarea de aprendizaje.