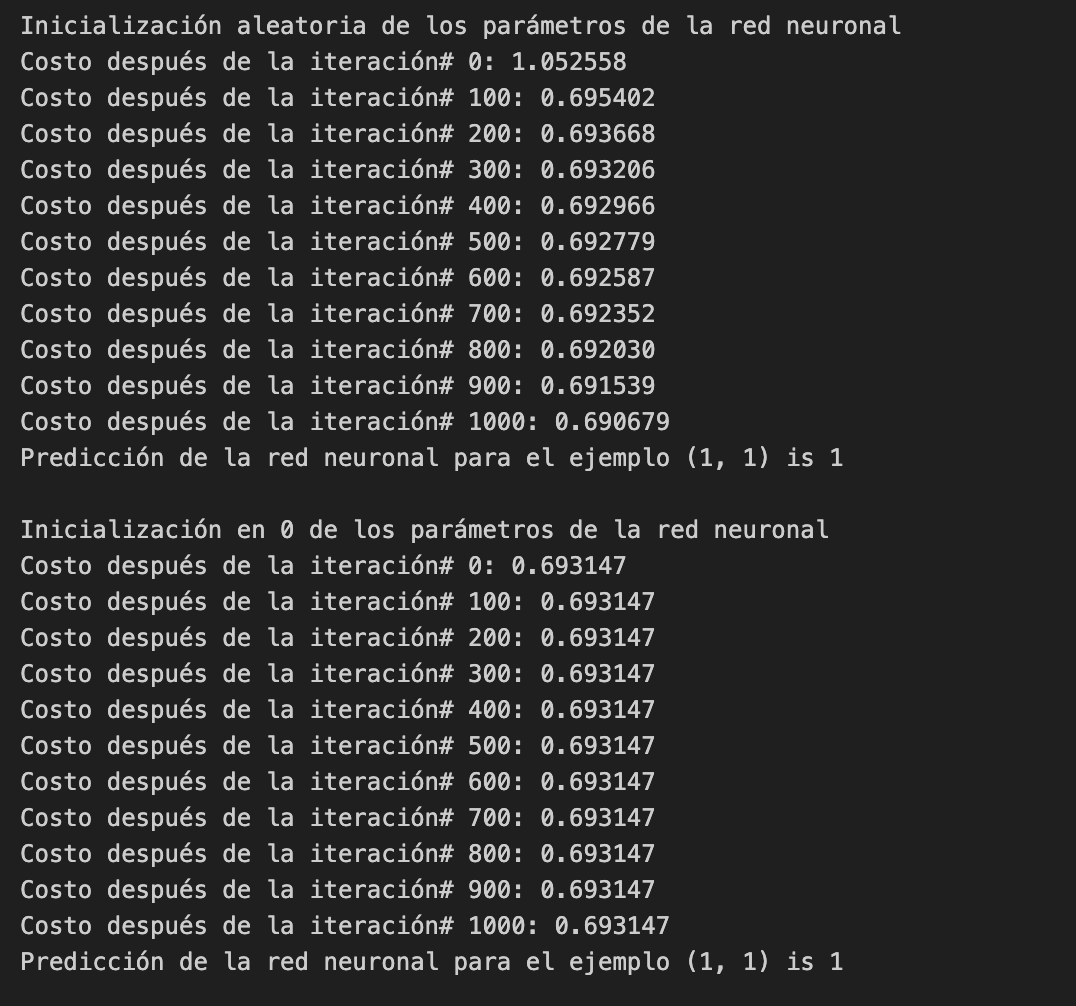
**Laboratorio 1**

1. *Existe diferencia entre la convergencia de los parámetros (pesos y sesgos) si estos son inicializados en 0 o como números aleatorios?*

**

Sí, existe una diferencia significativa. Si los parámetros se inicializan en cero, cada neurona en la red aprende de manera idéntica, lo que lleva a que todas las neuronas reciban las mismas actualizaciones durante el entrenamiento. Esto resulta en una simetría que impide que la red aprenda de manera efectiva, ya que no se produce una reducción significativa en la función de costo y los pesos y sesgos permanecen inmutables. Por otro lado, la inicialización aleatoria de los parámetros rompe esta simetría, permitiendo que las neuronas aprendan diferentes características desde el inicio. Esto generalmente lleva a una mejor representación del problema y a una convergencia más rápida, como se puede observar en la disminución gradual del costo a lo largo de las iteraciones.

1. *¿Qué diferencia en la convergencia de la función de costo y los parámetros existe si el learning rate del código es 0.01? 0.1? 0.5?*

*A screenshot of a computer

Description automatically generated*

El learning rate tiene un impacto significativo en la convergencia de la función de costo y de los parámetros. Con un learning rate de 0.01, la convergencia es muy lenta, lo que indica que la red está aprendiendo, pero el proceso es ineficiente. Con un learning rate de 0.1, la red converge de manera más eficiente y rápida, alcanzando un buen equilibrio entre la velocidad de aprendizaje y la estabilidad de las actualizaciones. Sin embargo, con un learning rate de 0.5, aunque la red converge rápidamente al principio, existe el riesgo de que el modelo salte alrededor del mínimo de la función de costo, lo que puede llevar a una convergencia menos estable y a predicciones inconsistentes. Por lo tanto, un learning rate intermedio como 0.1 suele ser más adecuado para lograr una buena convergencia sin comprometer la estabilidad del modelo.

1. *Implemente MSE como función de costo y propague los cambios en las funciones que lo requieran. ¿Qué cambios observa?*

*A screenshot of a computer

Description automatically generated*

Inicialmente, el costo con MSE puede ser diferente, pero lo más notable es el comportamiento de la convergencia. Con MSE, el costo disminuye de manera más gradual y estable sin grandes oscilaciones, lo que resulta en una convergencia más suave. Esto puede ser ventajoso para problemas de regresión donde se busca minimizar la diferencia promedio al cuadrado entre las predicciones y los valores reales. En general, la función de costo MSE proporciona una convergencia estable y una reducción constante del costo durante las iteraciones, manteniendo predicciones precisas y un comportamiento estable del modelo.

Link al repositorio: <https://github.com/mvrcentes/deep-learning_2024.git>