



Para resolver o dataset “Spiral”, utilizei apenas as entradas  $x_1$  e  $x_2$  e ajustei diferentes parâmetros da rede.

- Número de camadas: Redes muito rasas não conseguem representar o formato espiral. Com 2 camadas ocultas, o modelo já foi capaz de aprender o padrão.
- Neurônios por camada: Quantidades pequenas geraram subajuste. Com 8 neurônios por camada, a rede ganhou capacidade suficiente para separar as espirais.
- Ativação: A função Tanh apresentou melhor desempenho, pois produz fronteiras de decisão mais suaves para problemas altamente não lineares.
- Learning rate: Valores muito altos instabilizaram o treino. A taxa 0.03 proporcionou convergência estável e rápida.
- Regularização: A regularização L2 = 0.001 ajudou a suavizar a fronteira e evitou overfitting.
- Com esta configuração, o modelo atingiu Test Loss = 0.066, representando adequadamente o padrão espiral.