

Relatório de Experimento: Otimização em Redes Neurais (Dataset Spiral)

Disciplina: RNA e Deep Learning

Professor: Victor Casadei

Aluno: Gabriel Ferraz Milet

Turma: Engenharia e Análise de dados

Objetivo do Experimento:

Utilizar o TensorFlow Playground para projetar uma Arquitetura de Rede Neural capaz de resolver o Dataset Spiral, validando o impacto da profundidade, da largura e das funções de ativação.

Configurações Utilizadas:

Parâmetro	Valor Utilizado	Resultado Obtido
Dataset	Spiral (Não-Convexo)	Fronteira complexa totalmente separada.
Arquitetura (Camadas Ocultas)	4 Camadas	Profundidade necessária para construir as curvas da espiral.
Largura (Neurônios)	8 - 6 - 6 - 4	Largura moderada.
Ativação	ReLU	Escolhida por evitar saturação e garantir fluxo de gradiente estável.
Learning Rate	0.01 a 0.03	Taxa baixa, essencial para otimização estável em fronteiras delicadas.
Regularização (L2)	0.001 a 0.01 (Baixa)	Utilizada para estabilizar o treinamento (Não para evitar overfitting).
Test Loss	≈ 0.08 ou 0.09	Requisito de perda (< 0.1) atingido.

Resultados e Desempenho:

Resultados Finais: . Test Loss: $\mathbf{\backslash approx 0.08}$ (Atingido o requisito de perda < 0.1). Visualização (Screenshot): A rede conseguiu separar as classes azul e laranja, desenhando uma fronteira não-linear complexa que se "enrola" para seguir a forma da espiral.

O que Aprendi com o Experimento:

O experimento com o dataset Spiral foi fundamental para fixar o conceito de Hierarquia de Features e Não-Linearidade em Redes Neurais.

Profundidade Vence Largura: Para problemas com geometrias complexas e sobrepostas (não-convexos), o fator mais importante para o sucesso não é o número de neurônios por camada (largura), mas sim o número de camadas ocultas (profundidade). Cada camada adiciona uma nova transformação não-linear, permitindo que a rede construa a fronteira de decisão da espiral.

Ativação Crítica: A utilização da ReLU foi obrigatória. Redes profundas com Sigmoid ou TanH teriam sofrido com o desaparecimento do gradiente (vanishing gradient), impedindo o aprendizado eficaz das camadas iniciais.

Controle Fino de LR: O Learning Rate precisou ser mantido baixo (0.01 a 0.03). Uma taxa alta faria com que o modelo "pulasse" o ajuste fino das curvas da espiral (overshooting), impedindo que a Test Loss caísse abaixo do limite.

print do experimento:

