

Atividade Aula 4.3

Aluno: Guilherme Silva Câmara



Para atingir o objetivo de obter Test Loss menor que 0.1, foi necessário ajustar diversos hiperparâmetros da rede neural treinada no dataset Spiral. Para isso, foi utilizada uma configuração com três camadas escondidas, onde a primeira e a segunda continham 8 neurônios cada, e a terceira camada escondida continha 7 neurônios. A função de ativação escolhida foi a ReLU, devido à sua eficiência em problemas de classificação não linear. O learning rate foi definido como 0.01, o que proporcionou uma convergência estável ao longo de 1.000 épocas de treinamento. Além disso, foi aplicada uma regularização L2 com um fator de 0.001 para evitar o overfitting e melhorar a generalização do modelo.

O número de camadas teve impacto direto na capacidade do modelo, pois redes rasas não conseguem representar o formato espiral, enquanto três camadas forneceram profundidade suficiente sem tornar o modelo excessivamente complexo. A quantidade de neurônios também foi importante, pois valores intermediários permitiram boa expressividade sem provocar overfitting excessivo.

Utilizando a função de ativação ReLU, o treinamento foi acelerado e permitiu construir fronteiras complexas através da composição de camadas, mesmo sendo naturalmente menos suave que funções como tanh. O learning rate de 0.01 mostrou-se adequado para garantir convergência estável ao longo das 1.000 épocas. Já a

regularização L2 ajudou a controlar o crescimento dos pesos e reduzir o overfitting, mantendo o Test Loss próximo ao training sem prejudicar a capacidade de modelagem.

Por fim, a combinação desses hiperparâmetros permitiu resolver o problema de forma eficiente, mostrando como a combinação equilibrada de profundidade, número de neurônios, função de ativação, taxa de aprendizado e regularização influencia diretamente a qualidade da solução.