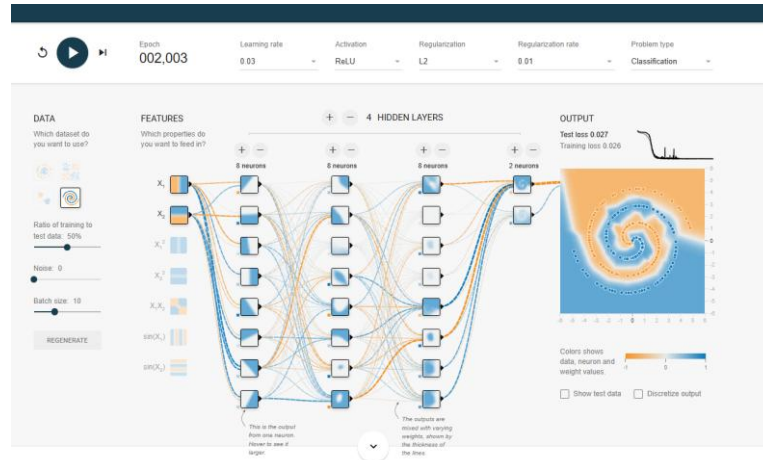


Atividade 4.3 – HCFN

Solução



**Comente sobre o que aprendeu nesse exercício (qual o impacto do número de camadas, neurônios em cada camada, ativação, learning rate e regularização).
Obs.: Test Loss: Abaixo de 0.1**

O número de camadas influencia diretamente a capacidade da rede de aprender padrões complexos. Redes com mais camadas conseguem capturar relações não lineares mais sofisticadas, enquanto redes com poucas camadas podem não separar adequadamente os dados do Spiral. No entanto, muitas camadas podem aumentar o tempo de treinamento.

O número de neurônios por camada também afeta a capacidade de modelagem. Mais neurônios permitem à rede aprender funções mais complexas, mas excesso pode causar overfitting. Por outro lado, poucos neurônios limitam o aprendizado, gerando underfitting.

O tipo de ativação define como os sinais passam entre as camadas. Funções não lineares, como ReLU, Sigmoid ou tanh, permitem que a rede capture padrões complexos. ReLU converge rapidamente, mas pode acabar ignorando alguns dados. Sigmoid e tanh são suaves e úteis para dados normalizados, mas podem saturar e tornar o aprendizado mais lento. Sem ativação não linear, a rede se comportaria como uma regressão linear, incapaz de separar curvas como o Spiral.

O learning rate controla a velocidade com que a rede ajusta seus pesos. Valores altos aceleram o aprendizado, mas podem causar oscilações e impedir a convergência. Valores baixos tornam o treino mais estável, porém mais lento.

Por fim, a regularização ajuda a evitar overfitting, penalizando pesos grandes ou desligando neurônios aleatoriamente. Sem regularização, a rede pode memorizar os dados de treino e performar mal no teste. Por outro lado, exagerar na regularização pode impedir que a rede aprenda padrões importantes, causando underfitting.