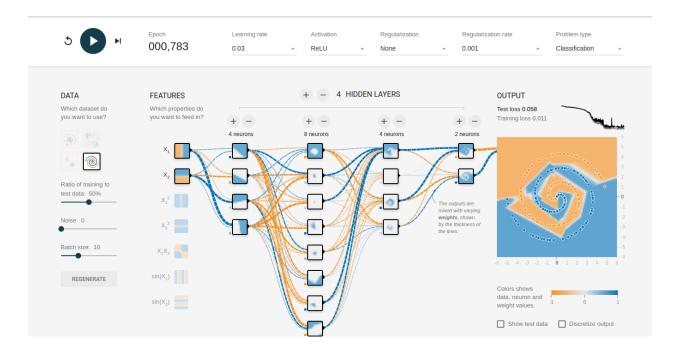
## Atividade Playground

Estudante: Israel Siqueira Email: imsj2@cesar.school



Os valores iniciais pré-estabelecidos pelo Playground foram os descritos abaixo. Com esses parâmetros, o modelo não deu sinais de que iria convergir para uma solução. Após aproximadamente 2 mil epochs, a perda havia estabilizado em 0.41 e 0.33 para os batches de Teste e Treinamento, respectivamente.

- Número de camadas: 2
- Número de neurônios em cada camada: 4 na primeira, 2 na segunda
- Tipo de ativação: tangente Hiperbólica (tanh)
- Learning rate: 0.03
- Fator de regularização: nenhum

Após algumas tentativas, o shape do meu modelo final foi:

- Número de camadas: 4
- Número de neurônios em cada camada: 4 na primeira, 8 na segunda, 4 na terceira e 2 na última
- Tipo de ativação: ReLu
- Learning rate: 0.03
- Fator de regularização: nenhum

Com essa rede obtive uma perda de teste de 0.058 e de treinamento 0.011 depois de aproximadamente 700 epochs.

Com o exercício foi possível observar a importância de escolher um Learning Rate adequado para a eficiência e momento do modelo. O LR de 0.03 e 0.01 se mostrou o mais adequado para essa configuração de rede. Mudanças no LR fizeram com que o modelo demorasse bem mais para convergir para um valor aceitável ou não convergir, mantendo uma perda alta mesmo após várias epocas.

A ativação ReLu também se mostrou superior às outras alternativas disponíveis. Usando-a fui capaz de obter resultados melhores e em menos epochs do que as demais.

Durante a prática também foi importante saber praticar o early stopping nos treinamentos, deixar o modelo rodando por várias épocas por vezes resultou em resultados com overfitting ou em que as perdas voltavam a aumentar após algum tempo, deixando o modelo impreciso. Ver a distribuição dos pesos entre os neurônios também foi interessante. No meu resultado final observei que alguns deles estavam contribuindo pouco para o output e que talvez poderiam ser removidos numa etapa posterior de otimização.