

Conclusões: Ao tentar diversas formas de utilização, a mais próxima do resultado desejado foi utilizando o Learning Rate de 0.03, a Ativação de Tanh, Regularization None e Regularization Rate 0.

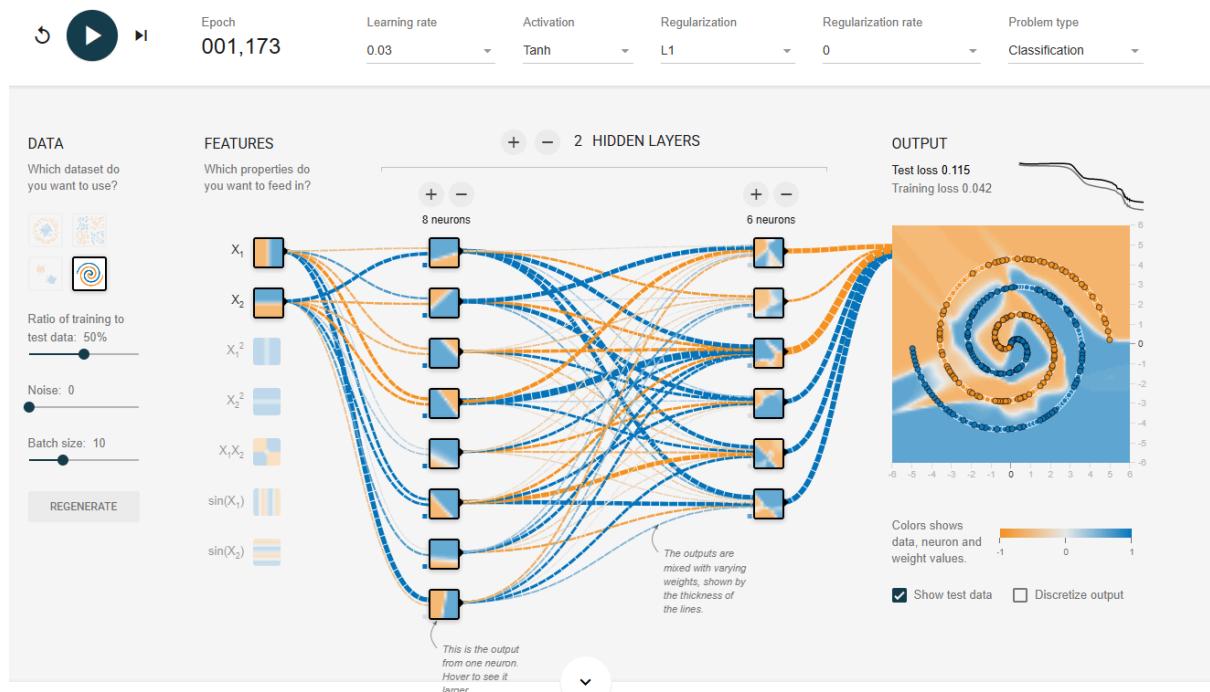
Foram adicionadas 2 Hidden Layers. A primeira com 8 neurons e a segunda com 6 neurons. o Test Loss cai consideravelmente ao adicionarmos mais neurons as Hidden Layers, bem como o número de épocas necessárias para chegar ao resultado desejado.

Ao realizar o teste, percebemos que aumentar a profundidade e a quantidade de neurônios melhora a capacidade do modelo de aprender o padrão de espiral. Isto pode acontecer por este ser um problema não linear.

A Ativação de Tanh traz resultados mais suaves e não lineares, o que casa bem com a solução do problema proposto.

O valor de Learning Rate de 1 encontramos dificuldade de convergência e muita oscilação, enquanto um valor mais baixo como 0.01 já tornava o modelo mais lento. Em 0.03 conseguimos chegar ao valor proposto em um número menor de épocas.

Outro ponto interessante é que com ou sem Regularization, conseguimos chegar ao resultado, porém utilizando Regularization L1, chegamos ao Test Loss 0.1 em 1.173 épocas, já com Regularization None chegamos ao resultado de Test Loss 0.1 em 1.684 épocas.



Epoch
001,684Learning rate
0.03Activation
TanhRegularization
NoneRegularization rate
0Problem type
Classification