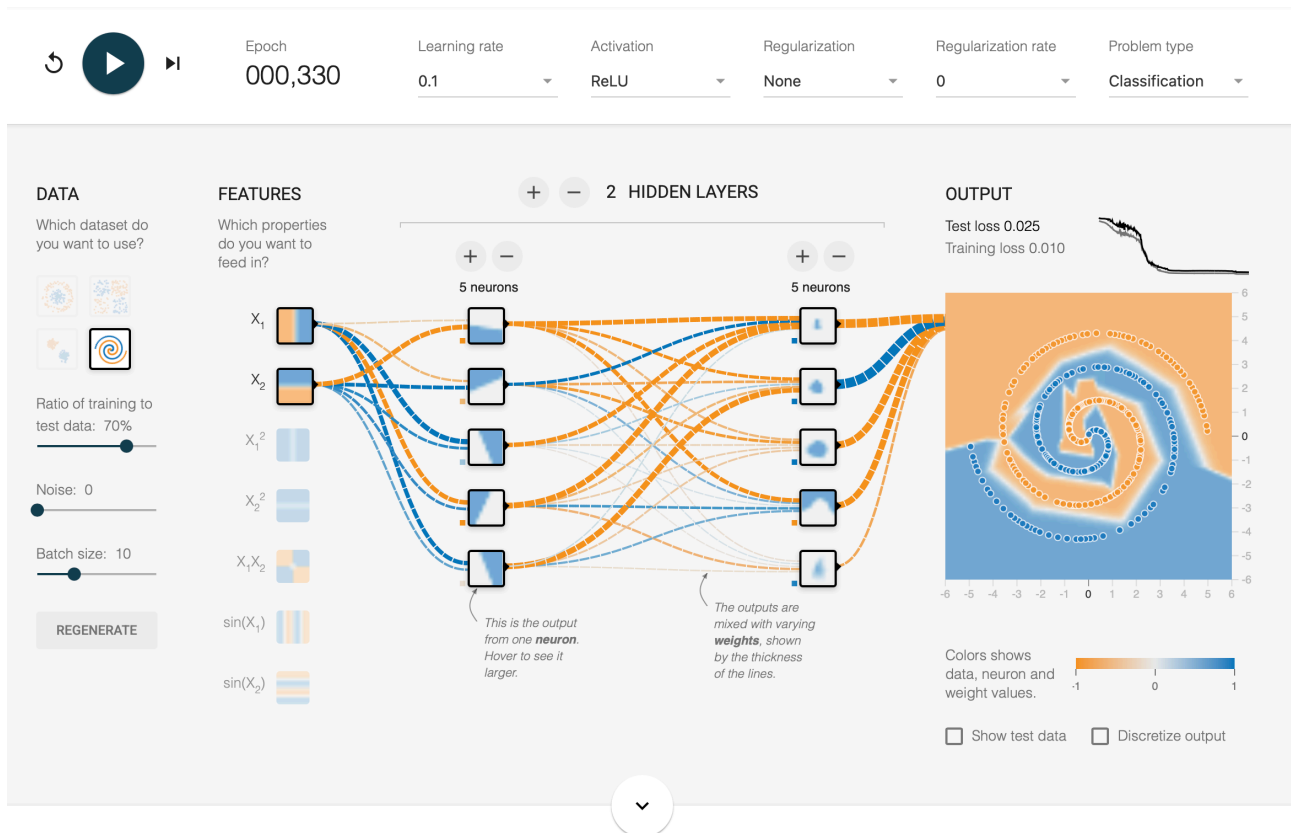


ATIVIDADE PLAYGROUND

Hélio Chaves Peixoto Neto
hcpn@cesar.school



Esse foi o menor número de camadas (2) e neurônios (10) usados para chegar a um *Test Loss* menor que 0.1 (nesse caso = 0.025, com 330 épocas) no *dataset* 'Spiral'

Em todos os contextos de uso, a Regularização não afetou positivamente o treinamento, seja a L1 ou L2 com uma taxa de 0.001 a 0.3.

Quando usei qualquer outro *Learning Rate* diferente de 0.1 e com outra função de ativação que não a ReLu, não obtive otimização do modelo para o valor de *Test Loss* pedido com essa quantidade de camadas e neurônios. Apesar disso, um *Learning Rate* de 0.03 me levou a um *Test Loss* = 0.108, mas foi preciso mais de 2 mil épocas.

O que mais ficou claro com essa atividade foi o fato das redes neurais serem uma caixa preta complexa de entender. Isso ficou evidente quando rodei o modelo com as mesmas configurações e o *Test Loss* foi 0.157, 6x maior que o obtido com a mesma configuração e o mesmo número de épocas, como mostrado no print abaixo. Ou seja, vários parâmetros têm influência no modelo, mas um primeiro passo para longe do mínimo global ou local pode fazer com que a otimização não seja possível a partir dali, mesmo utilizando as mesmas configurações.



Epoch
000,330

Learning rate
0.1

Activation
ReLU

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

Which dataset do you want to use?



Ratio of training to test data: 70%

Noise: 0

Batch size: 10

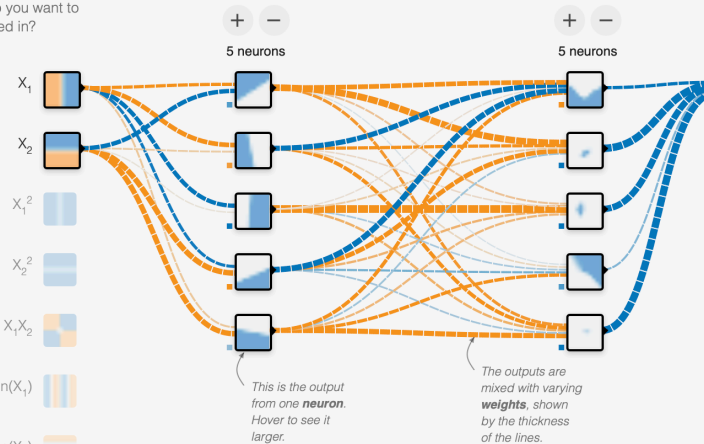
REGENERATE

FEATURES

Which properties do you want to feed in?

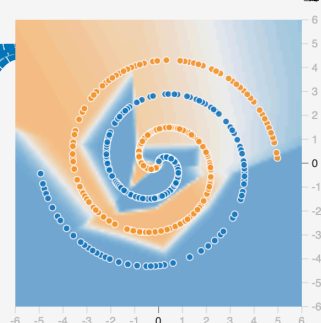
X_1
 X_2
 X_1^2
 X_2^2
 X_1X_2
 $\sin(X_1)$
 $\sin(X_2)$

2 HIDDEN LAYERS



OUTPUT

Test loss 0.157
Training loss 0.161



Colors shows data, neuron and weight values.

☐ Show test data ☐ Discretize output