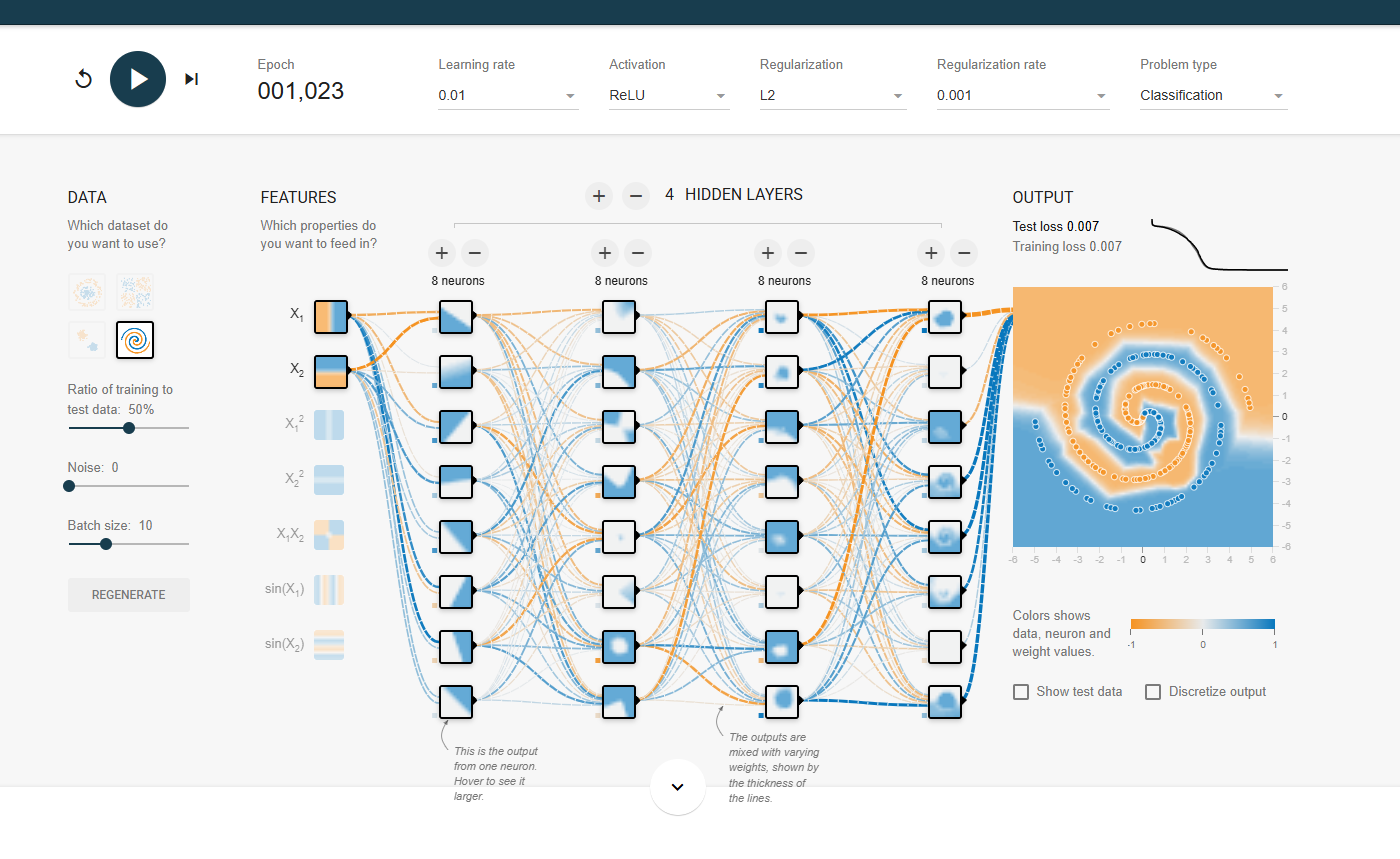
Pedro Henrique Costa Correia

phcc@cesar.school

Screenshot



Testei várias configurações até conseguir resolver o dataset Spiral. No final, a rede que funcionou melhor foi essa:

* 4 camadas com 8 neurônios cada
* Função de ativação: ReLU
* Learning rate: 0.01
* Regularização L2 com taxa 0.001
* Só usei x1 e x2 como entrada

Com isso, consegui deixar o test loss em 0.007, bem abaixo do 0.1 pedido. Percebi que quanto mais complexa a rede (mais camadas e neurônios), melhor ela entende padrões difíceis como o da espiral. Essa configuração chegou no resultado mais rápido.

Testei outros tipos de ativação, como a tanh, que até chega perto de 0.1, mas só depois de bastante tempo rodando. Com funções de ativação diferentes ou taxas de aprendizado mais altas, o modelo ficava instável ou não aprendia direito. No fim, essa combinação foi a que funcionou melhor.