

Atividade Playground

Aluno: Leandro de Andrade Moura

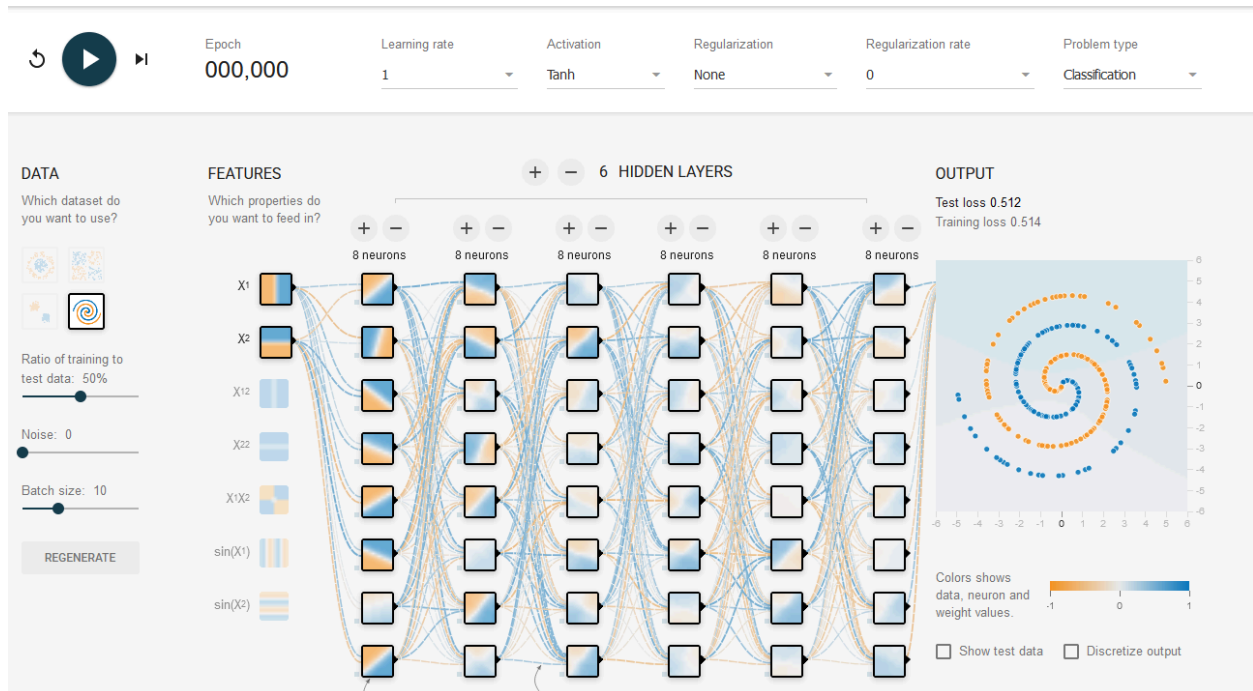
E-mail: lam5@cesar.school

Objetivo: resolver para o dataset "Spiral".

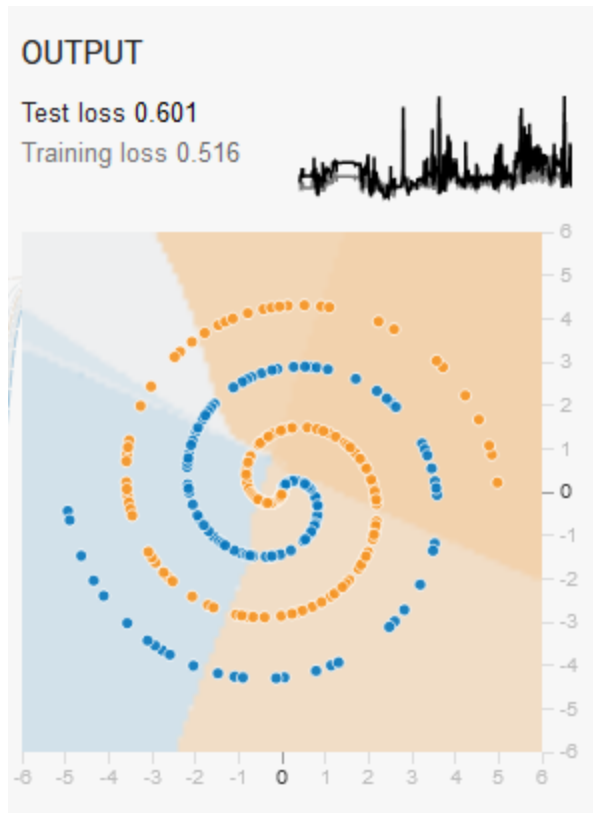
Restrições: utilizar apenas 2 entradas (x1 e x2).

Observação: buscar reduzir a quantidade de recursos e documentar o caminho percorrido.

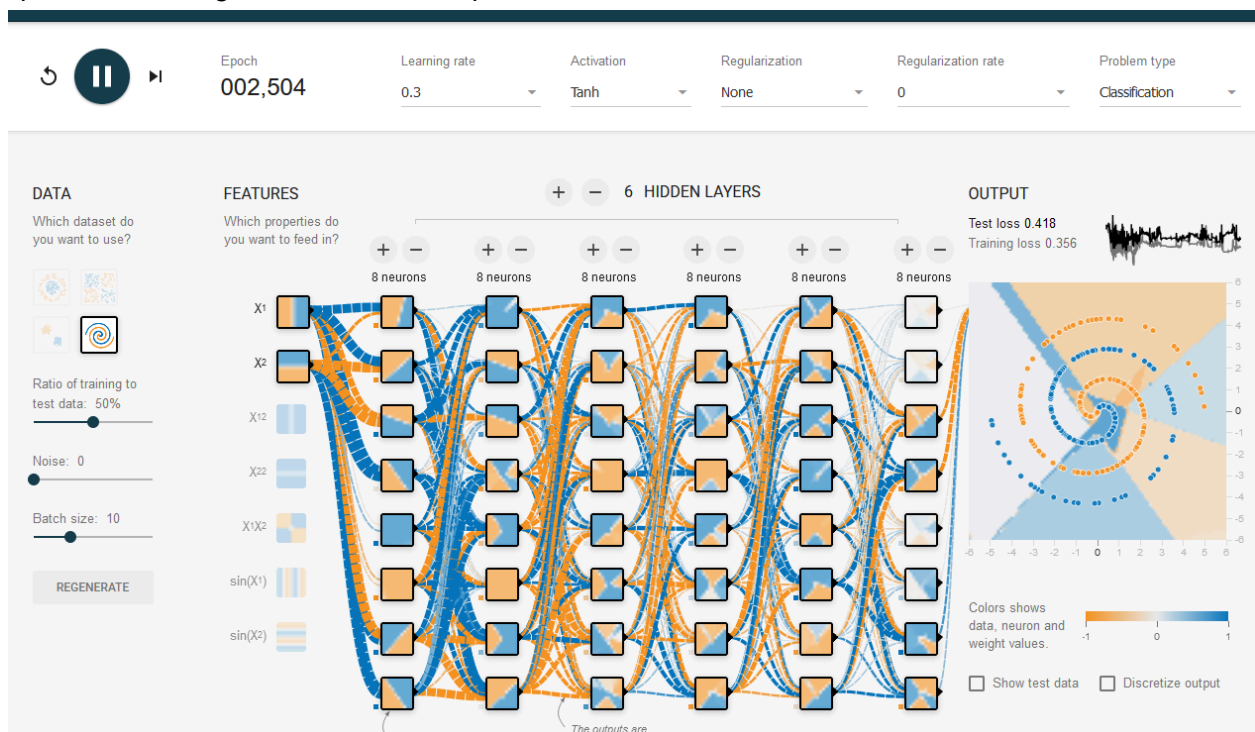
1. Tentando resolver com o máximo de recursos:



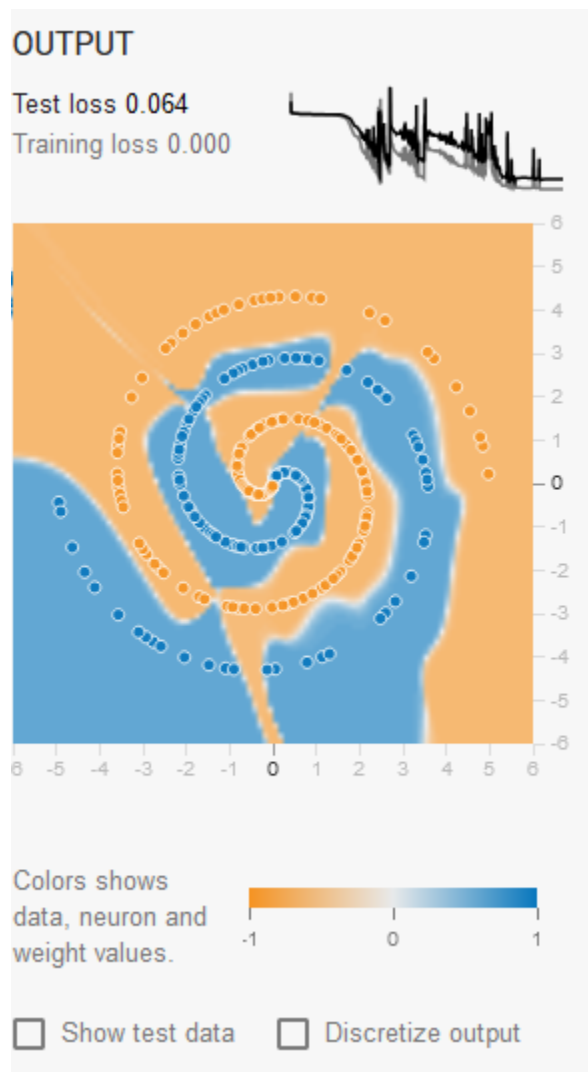
Iniciei mais agressivo com o *Learning Rate* maior que 1. A busca ficou instável. Setei para 1 e consegui obter avanços. No entanto, após 2 mil épocas, os resultados continuam a mostrar instabilidade.



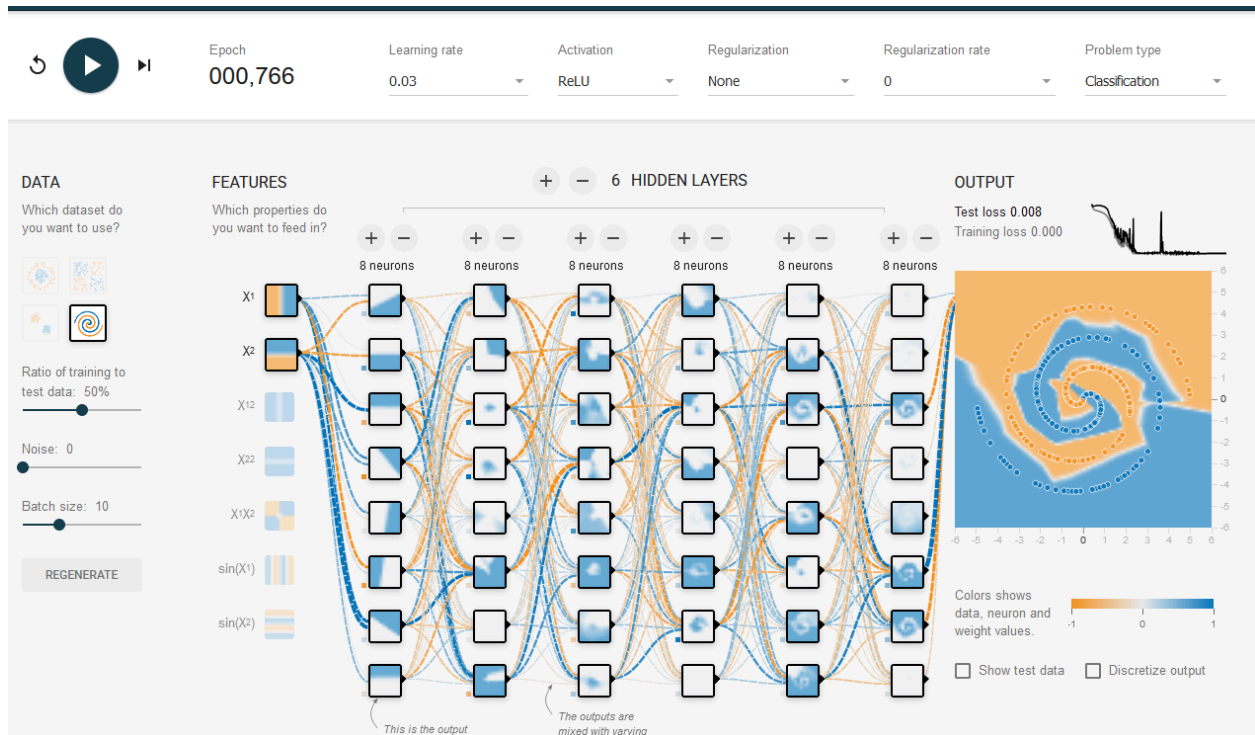
2. Resetei o processo para um *Learning Rate* de 0.3. Depois de 2500 épocas ele já tava apresentando algum desenho de espiral:



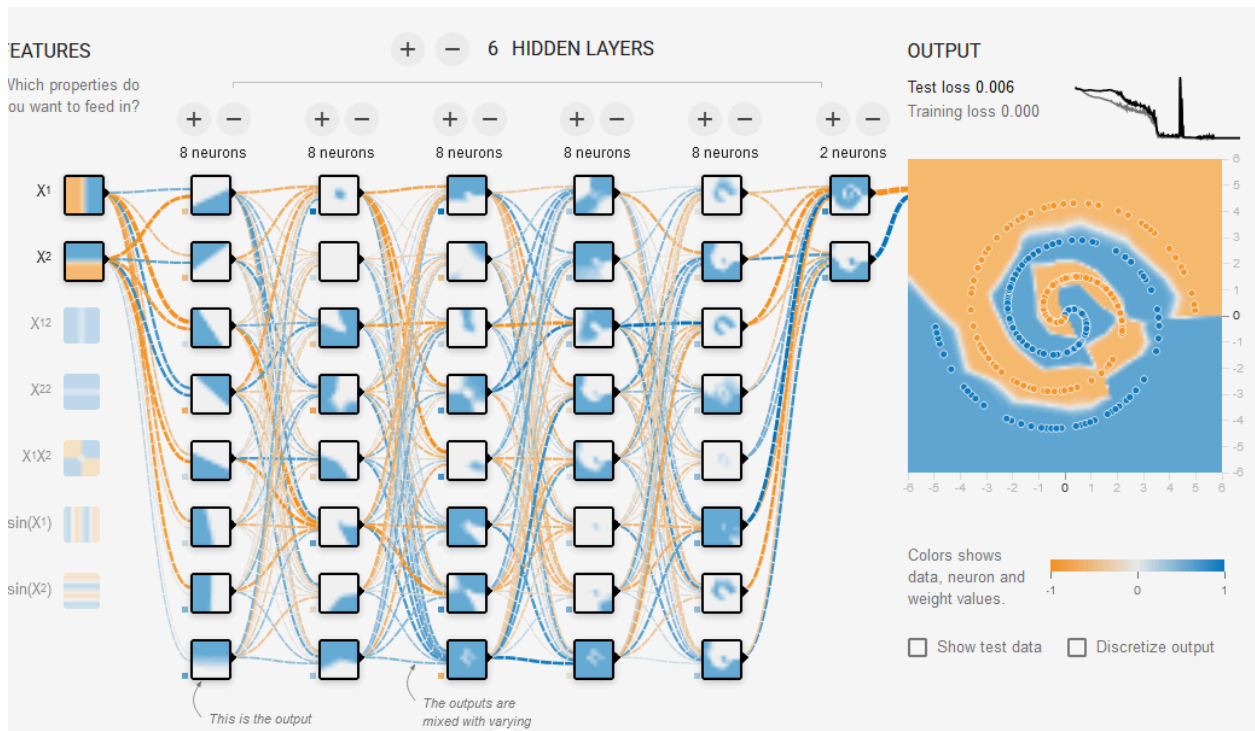
3. Ainda assim, durante o passo 2, houve muita flutuação na busca antes de alguma estabilidade. Reduzindo ainda mais o *Learning Rate* para 0.03. Após 1750 épocas, o *Training Loss* chegou a 0. No entanto, o *Test Loss* ainda não estava zerado. Ao explorar o *Test Data*, foi verificado que tinha uns pontos da espiral que o *Training Data* não cobriu (a região perto do zero do eixo X).



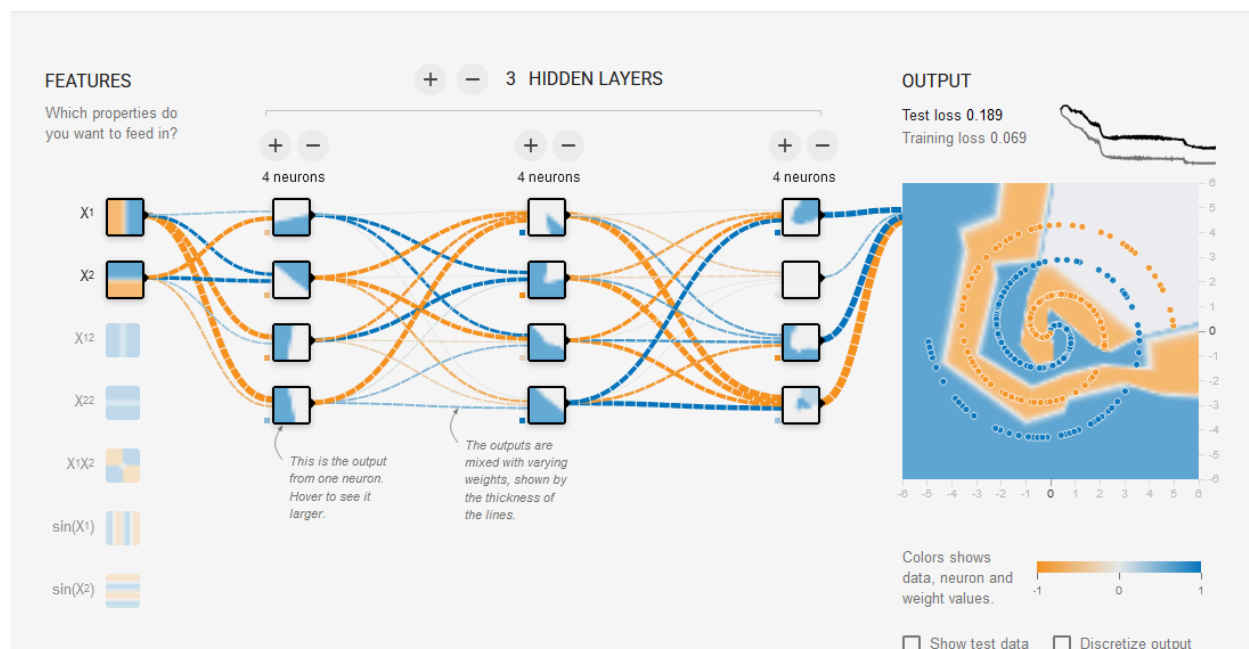
4. A função de ativação foi mudada para *ReLU*, observa-se que o modelo aprendeu muito mais rapidamente a desenhar contornos arredondados convergindo melhor. Em 750 épocas houve um erro bem pequeno no *Test Data*. Ao observar os neurônios, percebe-se que alguns ganharam a forma 'espiral' enquanto que outros ficaram em branco, o que sugere que o modelo é capaz de aprender esse formato em poucas camadas.



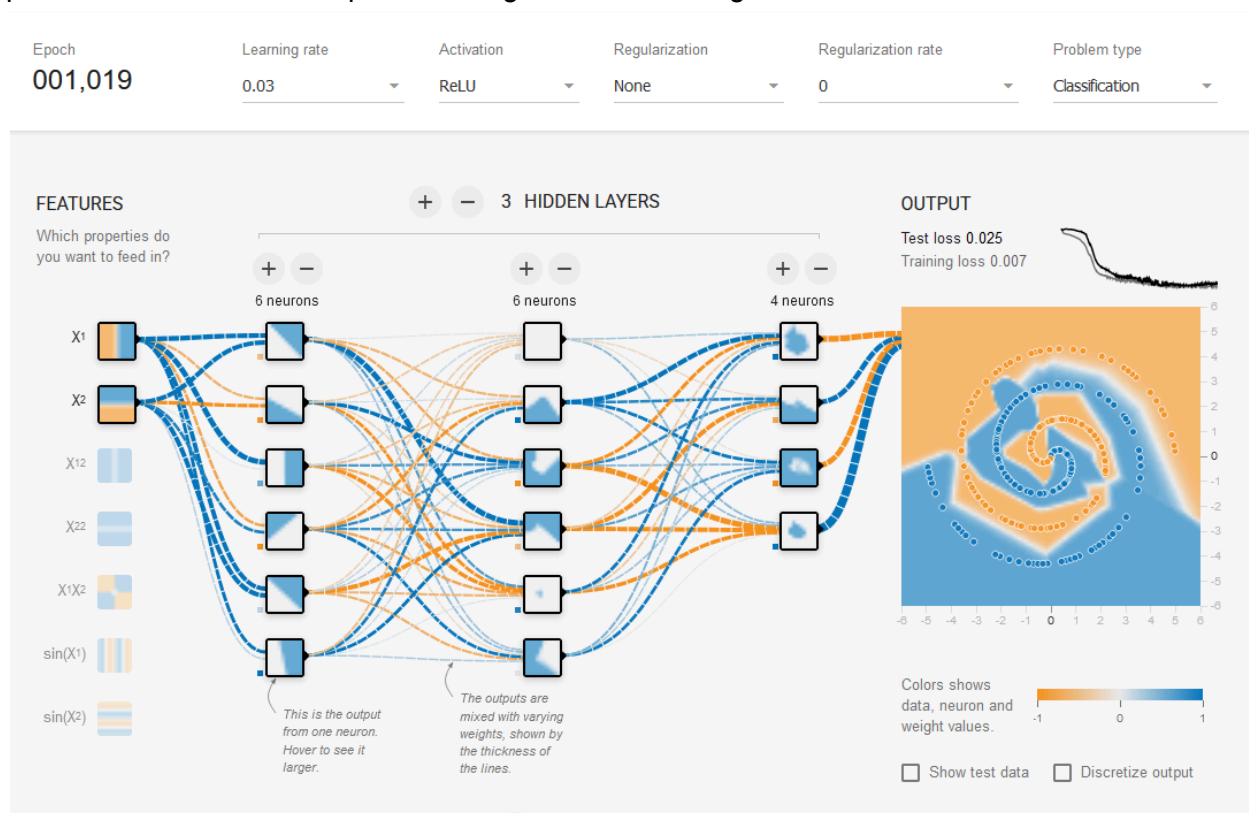
5. Com dois neurônios na última camada, foi possível chegar ao mesmo resultado. Uma olhada mais detalhada na imagem também sugere que há muitos neurônios em branco nas demais camadas, o que sugere uma forte capacidade de cortar neurônios.



6. Reduzindo a quantidade de neurônios pela metade. Neste ponto foi possível perceber que o segundo neurônio da última camada não teve papel nesse processo, mas o resultado ainda não foi o esperado. Isso sugere a necessidade de melhorar a complexidade no meio da rede.

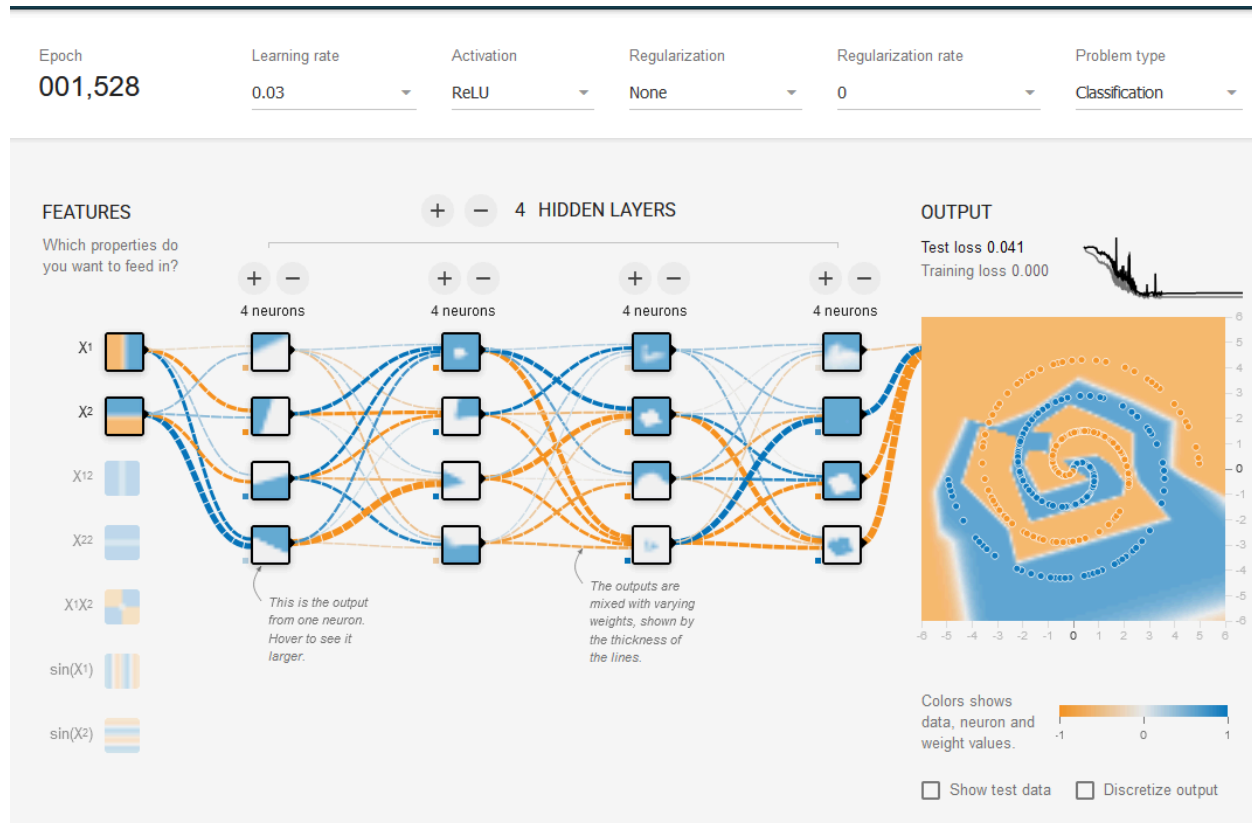


7. Após várias explorações, o aumento no número de neurônios nas primeiras camadas não parece traduzir corretamente a necessidade de complexidade do problema. Em alguns casos o problema teve dificuldade para convergir e em outros alguns neurônios não foram ativados.

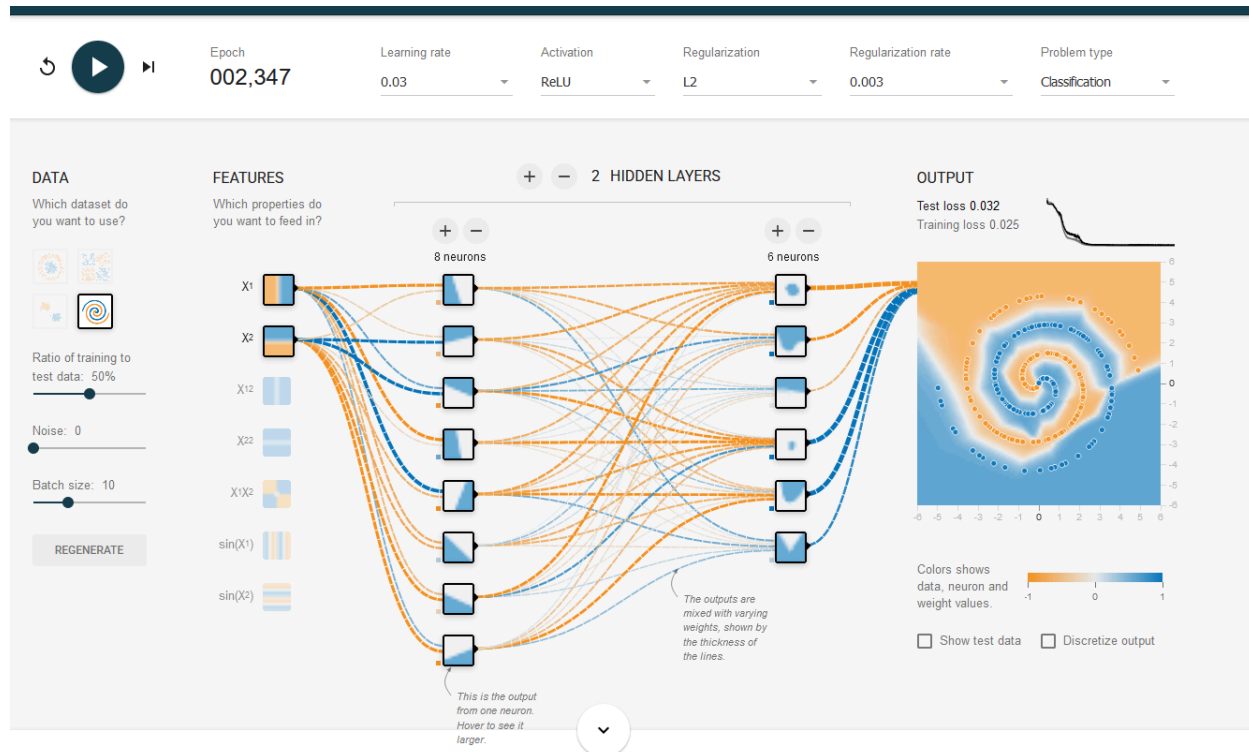


8. Após aumentar a quantidade de camadas, parece que o problema caiu num *Overfitting*, pois acertou todos os pontos, mas não generalizou bem. No entanto, na maioria dos casos, a rede

não conseguiu convergir. Parece estar tendo uma dependência bastante forte do estado inicial das camadas.

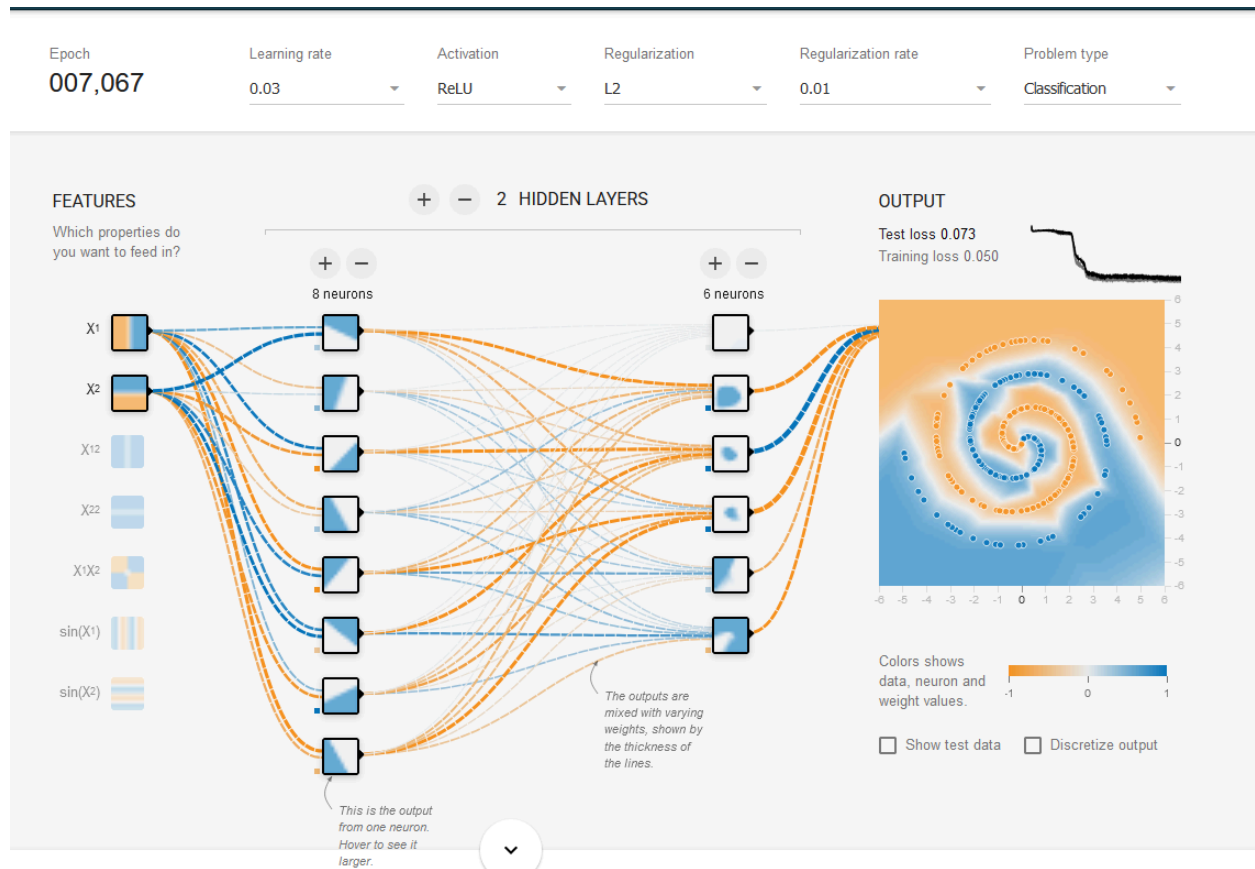


9. Diminui a quantidade de camadas, mas aumentei a quantidade de neurônio por camada. Durante os testes, também percebi que às vezes ficava preso em mínimos locais, então adicionei a regularização L2 com 0.03 de força. Ainda dá pra perceber que alguns neurônios não são tão ativados na segunda camada.



10. Resolvi aumentar ainda mais a força da regularização para tentar obter regiões mais 'fluídas'. Demorou bastante para convergir, mas o resultado pareceu ter sido ok. Notar que um dos neurônios mal foi ativado, o que sugere que dá pra diminuir ainda mais a quantidade de

neurônios.



11. Voltei para o passo 9. Achei que uma chance melhor seria tentar adicionar a regularização pra evitar o overfitting. Durante o treinamento também percebi uma certa instabilidade nas regiões de decisão, o que sugere um *Learning Rate* muito forte. Diminui o *Learning Rate*, mas entendi que isso iria aumentar ainda mais a quantidade de épocas. No entanto, o *Learning Rate* de 0.01 ainda estava forte demais. Removi a regularização e diminui ainda mais o *Learning Rate* para 0.003. No entanto, nenhuma dessas tentativas apresentaram resultados melhores que os anteriores, a saber, passo 8 e 10.