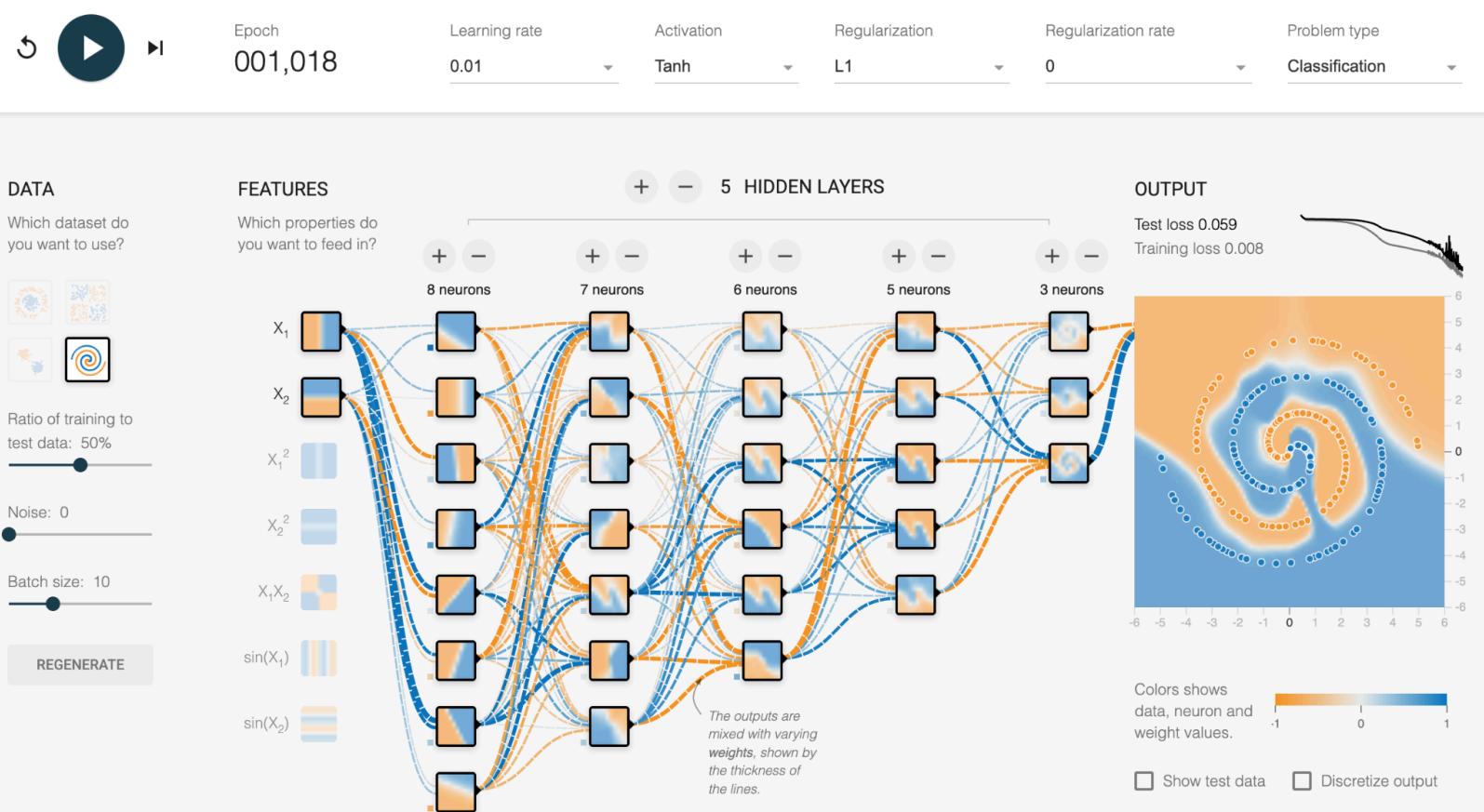


Especialização em Engenharia e Análise de Dados

RNA e Deep Learning

Atividade 4.3

Aluno: Marcel Pontes



A resolução da atividade de ajuste da rede neural para o problema de classificação espiral (Spiral) com duas variáveis de entrada e duas classes permitiu verificar o efeito das seguintes características do modelo ajustado na performance do mesmo em relação aos erros de treinamento e de teste.

- Número de neurônios e camadas

Resultados melhores, com relação ao tempo de treinamento e avaliação da classificação, foram obtidos com mais neurônios nas primeiras camadas (para processar features de ordem mais baixa) e menos nas subsequentes.

Além disso, um número de camadas mais próximo do máximo permitido pela plataforma auxiliou na formação de uma rede mais complexa que conseguisse separar uma distribuição não-convexa e espiralada usando apenas as variáveis de entrada em sua forma linear.

- Learning rate

O learning rate escolhido permitiu a convergência do algoritmo para um ponto satisfatório de perda, enquanto valores superiores testados causavam alterações tão altas nas atualizados dos pesos em cada época de treino que não favoreceram a convergência.

- Função de ativação

A escolha da tangente hiperbólica como função de ativação foi pelo fato de ter sido a que melhor favoreceu à característica curvilínea do dataset; porém, testes posteriores com a função ReLU possibilitaram uma convergência bastante rápida também. Por outro lado, outras funções desempenharam nada bem nessa tarefa.

- Algoritmo de regularização

A regularização L1 (Lasso), que aplica uma penalidade igual ao valor absoluto dos coeficientes do modelo à função de perda, favorecendo matrizes mais esparsas, levou a significativas oscilações no treinamento, mas auxilia na mitigação do *overfitting* e proporciona um erro no conjunto de teste tão baixo como foi visto.