

Nome: Cassiane Bueno  
E-mail: [cfsb@cesar.school](mailto:cfsb@cesar.school)



Dentre as várias simulações que fiz, percebi que a ativação ReLU era mais indicada para o tipo de dataset (espiral). Fixando a ativação, a learning rate e a regularização e modificando o número de camadas e de neurônios, percebi que:

- não necessariamente o aumento de camadas e de neurônio gera menor perda com menos épocas;
- o aumento do número de camadas gerou mais estabilidade;
- após certo número de épocas o valor da perda volta a aumentar e oscila, gerando instabilidade;
- muitas camadas e muitos neurônios gerou overfitting.

Por exemplo, seguem algumas das simulações que fiz com ativação ReLU, learning rate de 0,03 e a regularização igual a None:

- **3 camadas de 6, 5 e 6 neurônios a perda chegou a 0.065 com 352 épocas. Essa foi a da imagem acima escolhida por ter menor perda com menos épocas, menos camadas e menos neurônios.**
- 4 camadas de 6, 5, 5 e 6 neurônios a perda chegou a 0.091 com 473 épocas.
- 4 camadas de 6, 4, 4 e 6 neurônios a perda chegou a 0.098 com 559 épocas.
- **4 camadas de 8, 6, 6 e 8 neurônios a perda chegou a 0.016 com 371 épocas. Menor perda obtida, porém com mais camadas e mais neurônios.**

Ao alterar a learning rate também percebi que:

- diminuir a learning rate demora mais para ajustar os pesos;
- aumentar a learning rate também não significa que melhora a acurácia e eficiência, fiz as simulações com maior learning rate de 0.1 e os resultados não foram melhores. Por exemplo, para o caso impresso, precisou de 514 épocas para chegar a uma perda de 0.052.