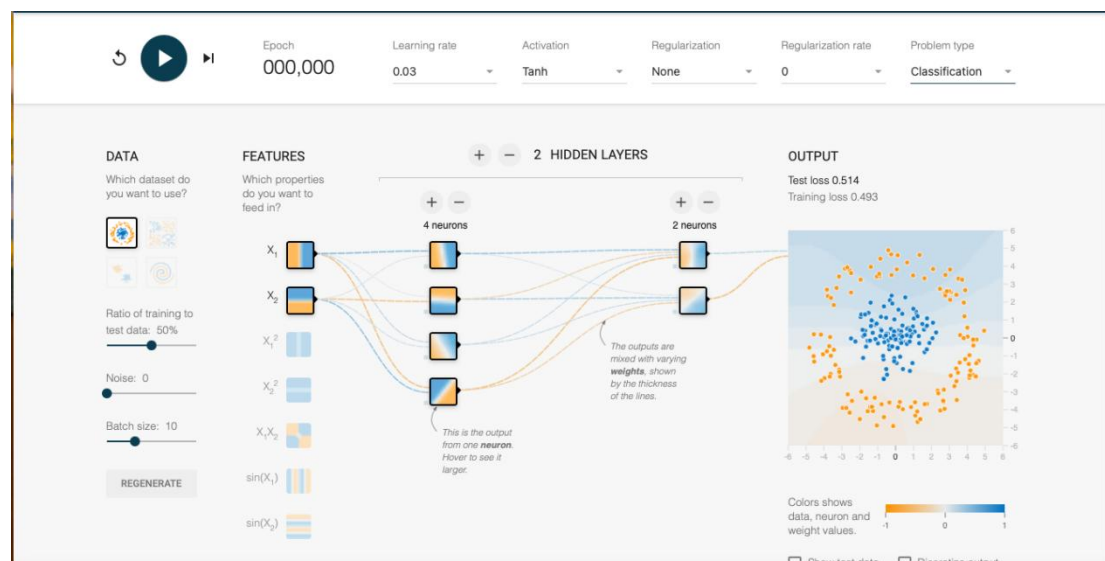


Tensorflow playground 试用报告

122106222784 贺梦瑶

1、页面构成

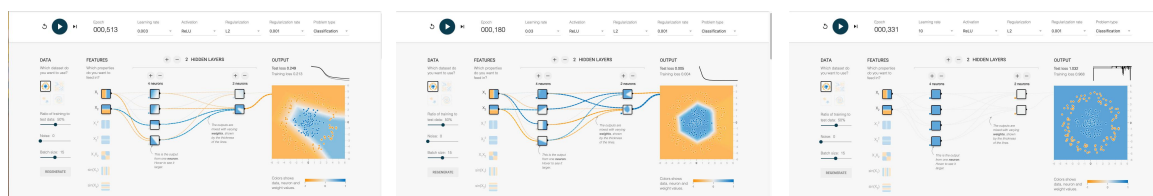


如图，页面上方为一列控制神经网络的参数，从左到右分别是，训练的开关、迭代次数、学习速率、激活函数、正则化、正则化率 and 问题的类型。下方为主要分为 DATA（数据），FEATURES（特征），HIDDEN LAYERS（隐含层），OUTPUT（输出层）。DATA 一栏里提供了 4 种不同形态的数据，分别是圆形、异或、高斯和螺旋。平面内的数据分为蓝色和黄色两类，可以调节噪声、训练数据和测试数据的比例和 Batch size 的大小。FEATURES 一栏包含了可供选择的 7 种特征： X_1 （横坐标分布的数据特征）、 X_2 （纵坐标分布的数据特征）、 X_1^2 （非负的抛物线分布）、 X_2^2 （非负的抛物线分布）、 X_1X_2 （双曲抛物面分布）、 $\sin(X_1)$ 、 $\sin(X_2)$ （正弦分布），通过这些特征的分布组合将两类数据区分开。HIDDEN LAYERS 一栏可设置多少隐含层。OUTPUT 一栏将输出的训练过程直接可视化。

2、试用

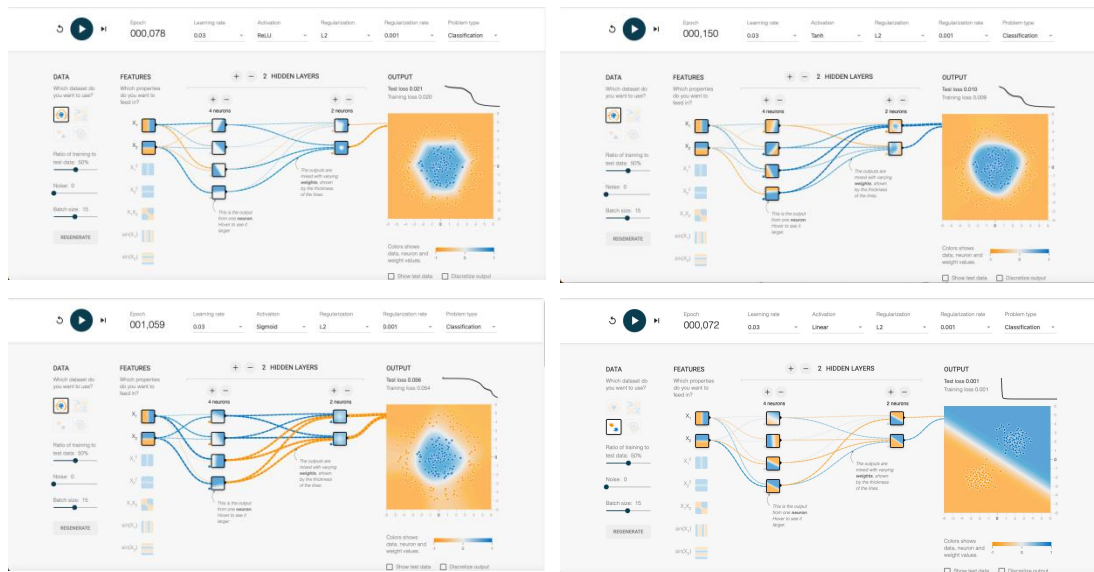
(1) 学习率

学习率越大，输出误差对参数的影响就越大，参数更新的就越快，但同时受到异常数据的影响也就越大，很容易发散。学习率越小，则会导致更新速度慢，影响效率。如图，学习率 0.003 导致更新过慢，学习率 10 导致学习结果无法稳定。

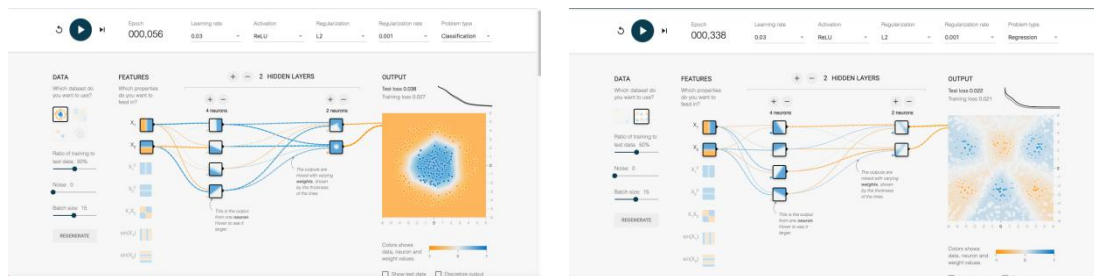


(2) 激活函数

选择 Sigmoid 函数作为激活函数，发现训练的时间很长，ReLU 和 tanh 函数能大大加快收敛速度，Linear 函数仅适用于能线性分割的情况。



(3) 任务



(4) 隐藏层个数

选用了3层隐含层，每层特征个数为8,8,2的模型和6层隐含层，每层特征个数为8,8,8,8,8,2的模型。3层隐含层模型大概200多步就达到了test loss为0.013，training loss为0.001，而6层隐含层模型跑了500多步，test loss为0.016，training loss为0.000，有点过拟合。

