TensorFlow Playground试用报告

TensorFlow Playground 是由 Google 开发的一款基于 Web 的交互式工具，是一个用来图形化教学的简单神经网络在线演示和实验的平台，非常强大且极其易用。它提供了一个用户友好的界面，允许用户实验和可视化人工神经网络的行为，用户可以调整各种网络参数，比如层数、每层神经元数量、激活函数和学习率等。通过修改这些参数，用户可以看到网络的反应，并学习不同配置对网络性能的影响。这个工具可以帮助用户更好地理解人工神经网络的工作原理。

1. TensorFlow Playground功能简介

TensorFlow Playground的布局划分如图1.1所示，总体上有如下区域：

* 运行控制区：这里主要对算法执行进行控制，可以进行启动、暂停和重置操作；
* 迭代次数展示区：这里展示当前算法执行到了哪一次迭代；
* 超参数选择区：这里可以调整算法的一些超参数，不同的超参能解决不同的算法问题，得到不同的效果；
* 数据集调整区：数据集定义了我们要解决怎样的问题，数据集是机器学习最为重要的一环；
* 特征向量选择：从数据集中选取出可以用来被训练的特征值；
* 神经网络区域：算法工程师构建的用于数据拟合的网络；
* 预测结果区，展示此算法的预测结果。

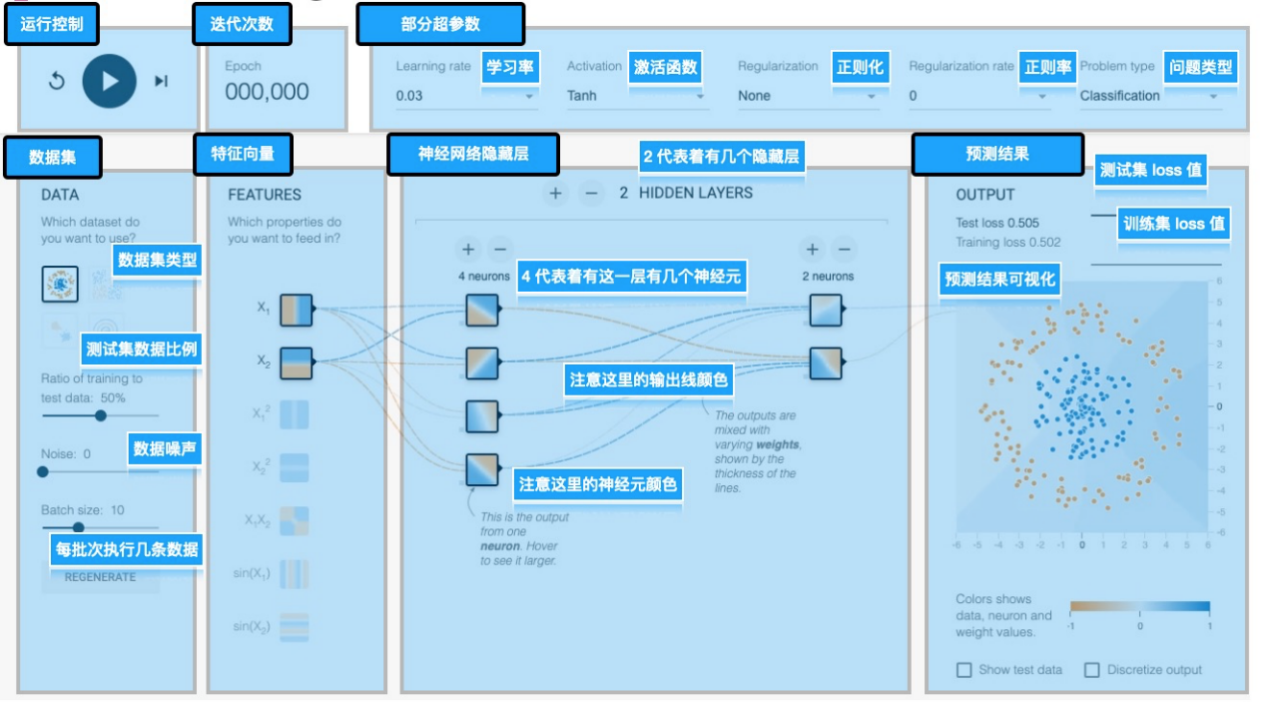


图1. 1 TensorFlow Playground区域划分

如图1.2所示，TensorFlow Playground提供了4种不同形态的数据。每组数据，都是不同形态分布的一群点，每个点都有两个特征（x1和x2），表示点的位置，并且可以被归为两个不同的类别：橙色和蓝色。这两个类别分别可以用圆形、两条正交直线、对角线和螺旋线的函数来划分。神经网络的目标就是通过训练，知道哪些位置的点是橙色、哪些位置的点是蓝色。

相较于前三个数据集，最后一个数据集函数拟合最困难，是TensorFlow Playground 里调试起来最有难度的数据集。

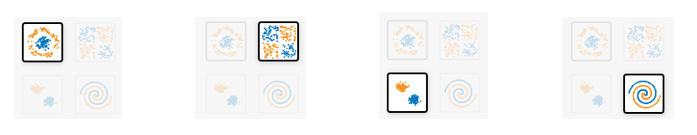


图1. 2 数据可视化

TensorFlow Playground中的数据十分灵活。可以调整noise（干扰）的大小，还可以改变训练数据和测试数据的比例多少。下图是不同noise的数据分布：

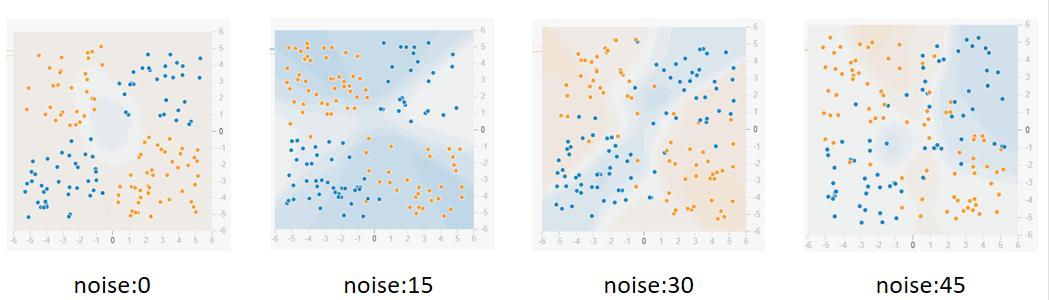


图1. 3 不同noise数据分布

TensorFlow Playground中的数据只有二维特征，这在处理任务时效果可能不好，所以它提供了由这两个特征衍生出的其他7种特征，单纯的特征值大小，特征值的平方，和特征值的正弦分布。如图1.4所示。这些特征可以随意组合进行特征选择作为模型的输入。

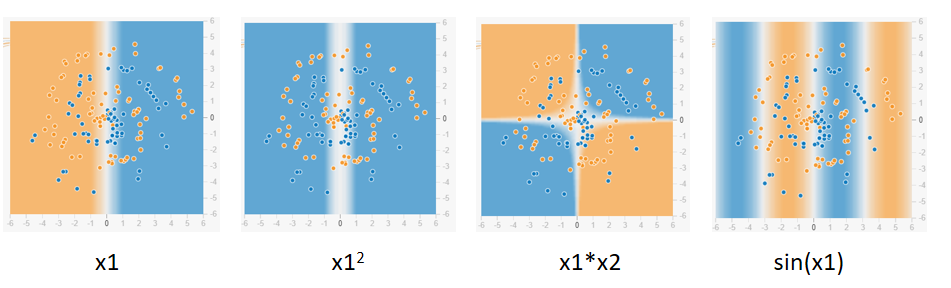


图1. 4 数据特征

TensorFlow Playground可以动态修改模型的隐藏层层数和每层的单元个数，针对同一任务，我们可以进行模型结构的调整使得结果达到最优。不同的隐藏层如图1.5所示：

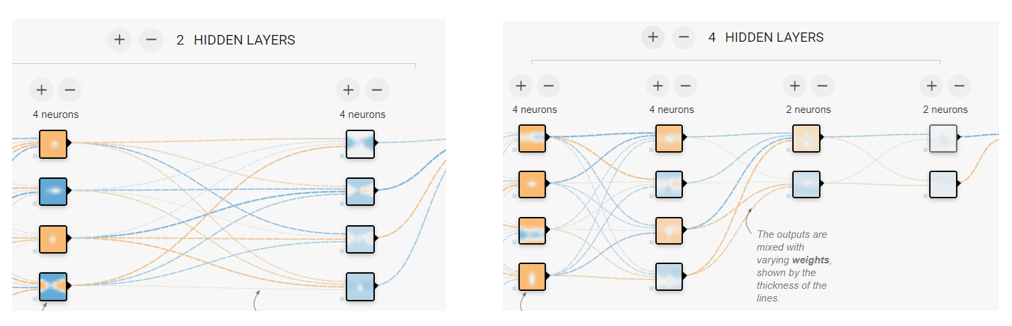


图1. 5 隐藏层动态调整

在选择完数据，完成特征工程，并调整模型结构后，最后需要在超参数选择区调整超参数进行模型训练，例子如图1.6所示。

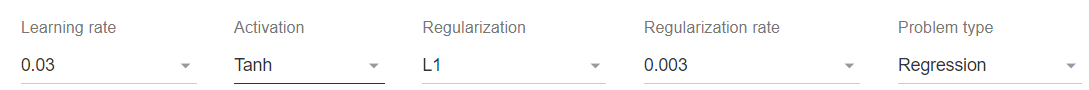


图1. 6 超参数选择

1. 试用样例

* 样例1，参数设置：学习率0.0001，激活函数Tanh，正则项L1，比例0.003，问题类型Regression。数据：选择第一个。网络结构：选择含两个隐藏层，第一个隐藏层4个神经元，第二个隐藏层2个神经元。可以看出几乎完美地分离出了橙色点和蓝色点。

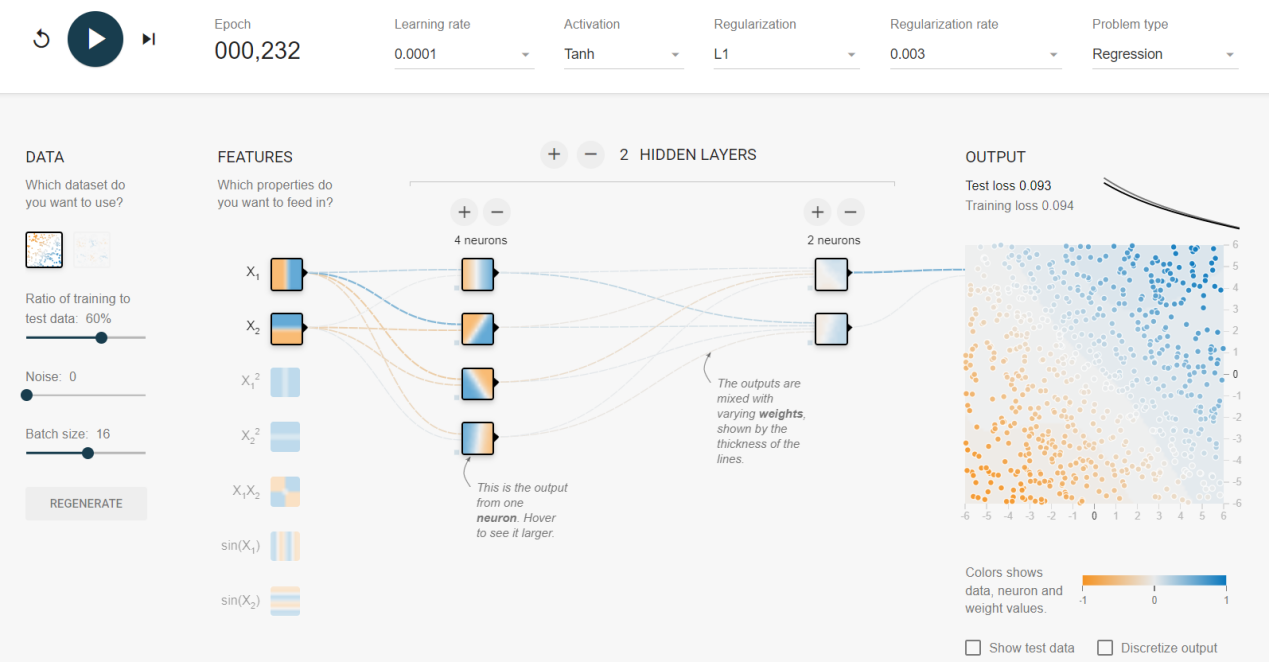


图2. 1 样例1

* 样例2，参数设置：学习率0.0001，激活函数Tanh，正则项L1，比例0.003，问题类型Classification。数据：选择左上第一个。网络结构：单个隐藏层8个神经元，可以看出单个隐藏层增加神经元也可以处理的很好。

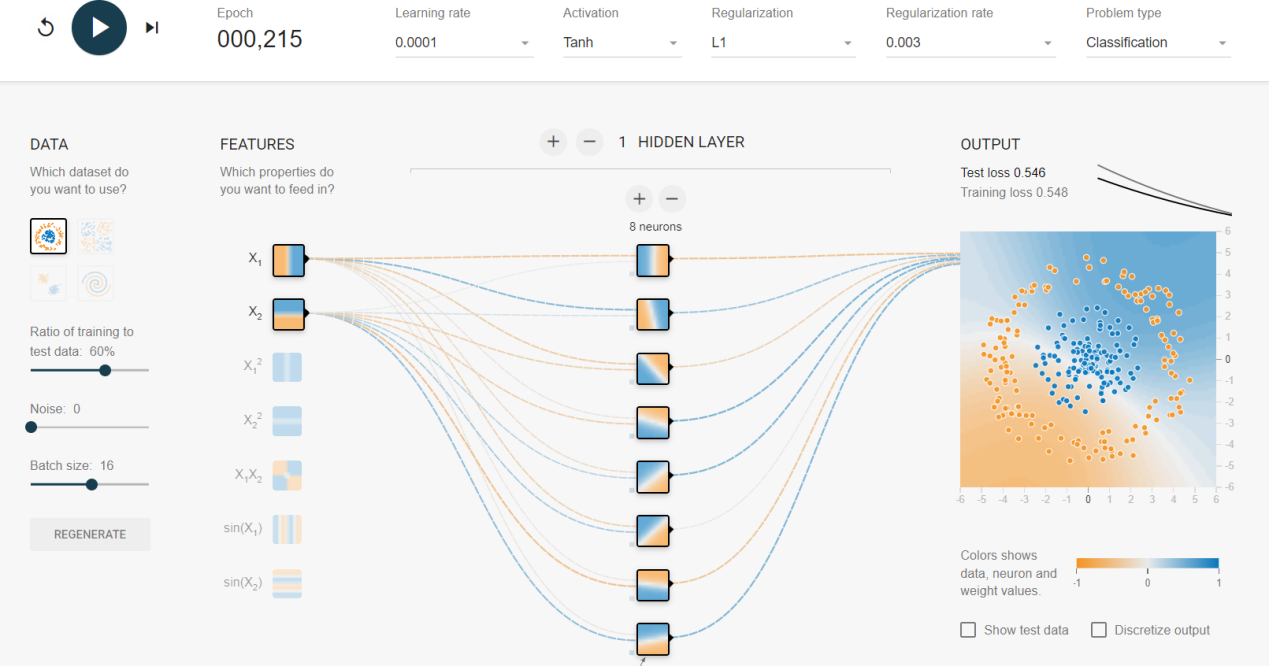


图2. 2 样例2

* 样例3，参数设置：学习率0.0001，激活函数Tanh，正则项L1，比例0.003，问题类型Classification。数据：选择右下角。网络结构：选择含6个隐藏层，各有8个神经元。只需要输入最基本的特征x1, x2, 只要给予足够多层的神经网络和神经元，神经网络会自己组合出最有用的特征。在难度最大的数据集上也完美的区分了不同的数据点。

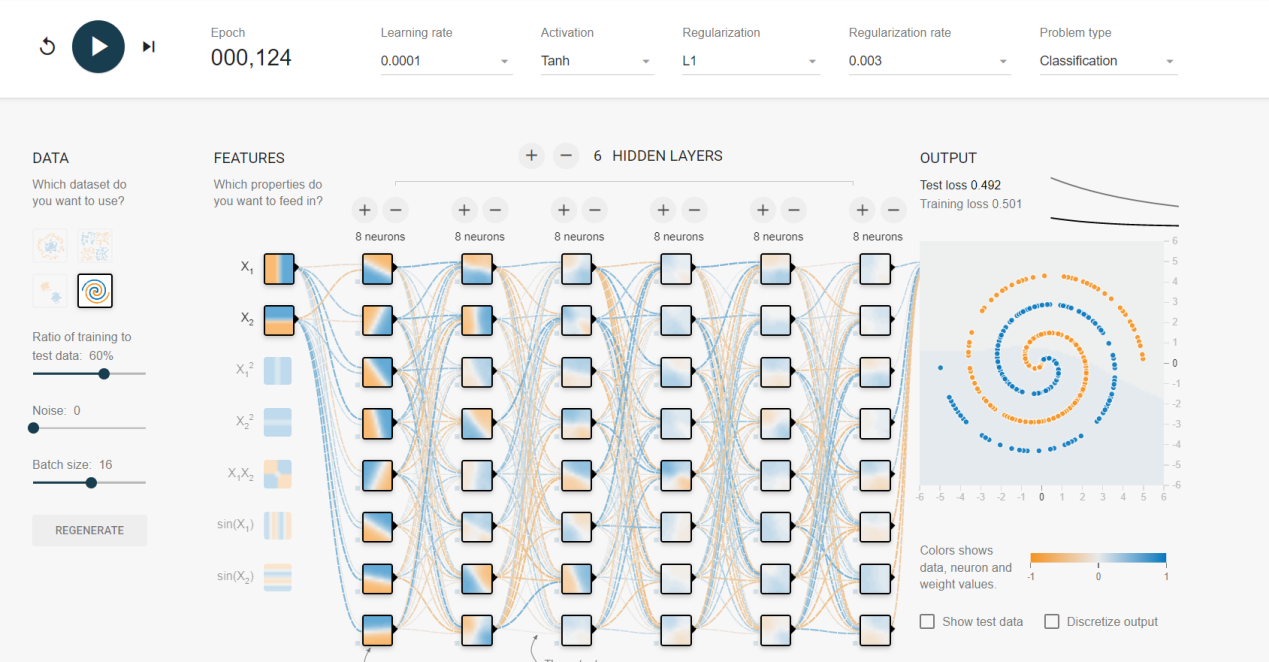


图2. 3 样例3

1. 总结

TensorFlow Playground具有直观的界面，能够显示神经网络和训练数据的图形化表示。用户可以通过拖拽、点击等方式，调整网络的各种参数，例如层数、每层神经元的数量、激活函数和学习率等。这些参数的修改可以直接反映在界面上，帮助用户更好地理解神经网络的行为和性能。

其次，TensorFlow Playground提供了内置的数据集和可视化工具，使用户可以更加深入地了解神经网络的工作原理。这有助于用户理解网络是如何处理输入数据的，从而优化网络的性能。

总之，TensorFlow Playground是一款非常有用的工具，它为用户提供了一个直观、易于使用的界面，允许他们调整神经网络的各种参数，并实时可视化网络的行为和性能。它可以帮助用户更好地理解神经网络的工作原理，并通过实验找到最佳的网络配置，从而优化神经网络的性能。