

对比 GAN、WGAN、WGAN-GP

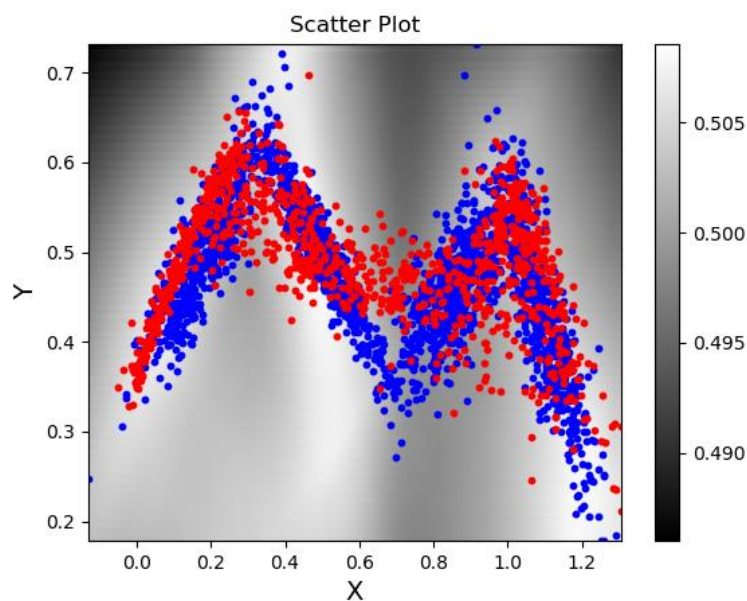
实验中，首先将样本随机打乱顺序，然后取其中的 6000 个点进行训练；剩余的 2192 个点用来画分布图，来验证生成的数据是否拟合该分布。

在训练完网络后，利用生成网络生成 1000 个点，画在图中，检查是否符合测试集中点的分布。测试集的点用蓝色画出，生成的点用红色画出。

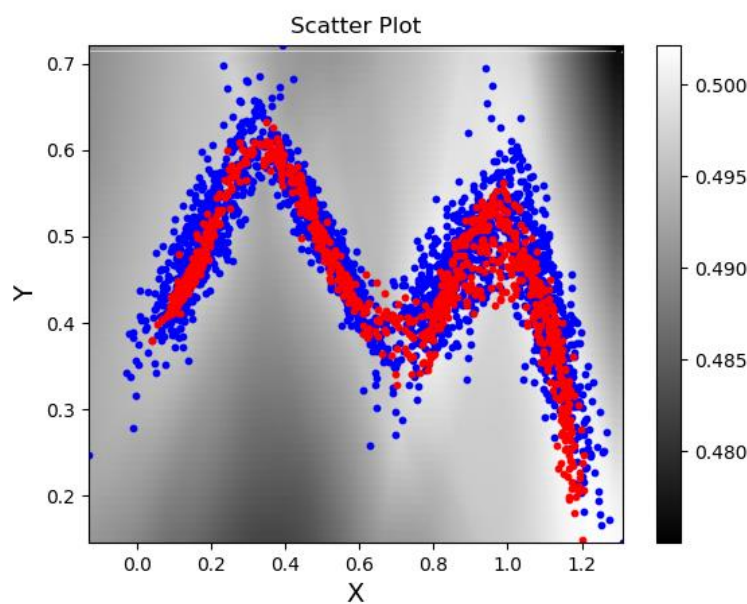
参数设置：BatchSize 设成了 150；三个网络分别遍历了 100 次训练集；LR = 0.0003

GAN:

原始 GAN 的效果比较好，第 20 个 epoch 已经可以看到大概的分布了：



在训练完 100 个 epoch 后，生成的点分布如下：



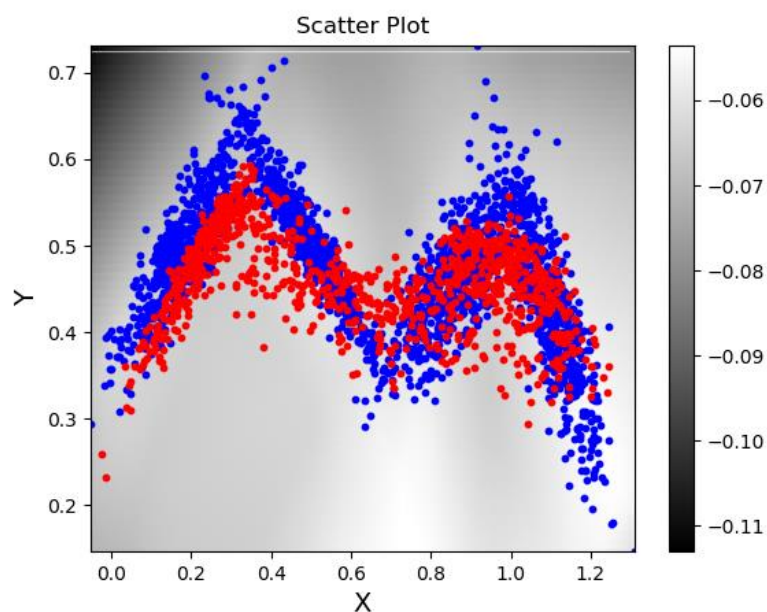
WGAN:

WGAN 相比于 GAN 主要有四点改进:

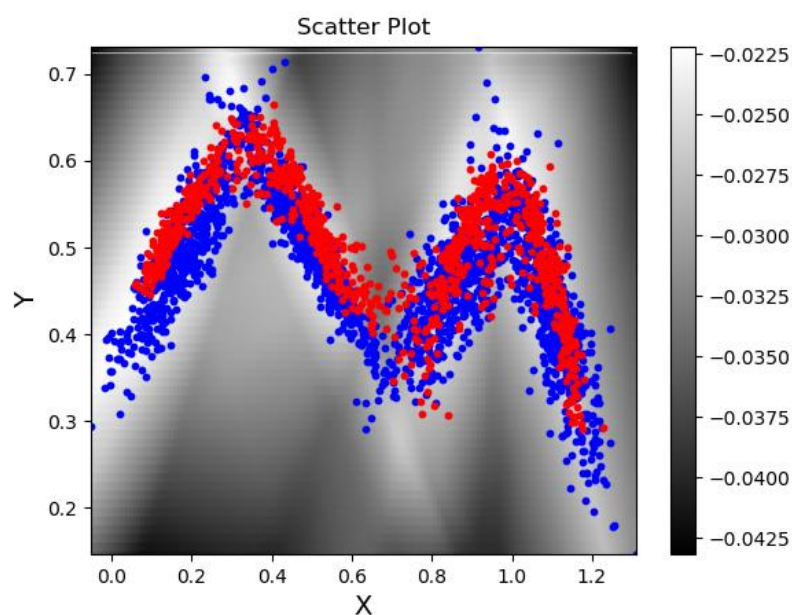
1. discriminator 的最后一层取消 sigmoid
2. discriminator 的 w 取值限制在 $[-c, c]$ 区间内, 确保 lipschitz 连续
3. 改为非 log 的 loss
4. 不使用具有动量的优化方法, 比如使用 Adam, 转而使用诸如 RMSProp, SGD 等方法 (gan 已经实现, 不需要改了)

具体实现在代码中。

WGAN 的效果比 GAN 要差一些, 第 40 个 epoch 才出现 M 型分布:



到 100 从 epoch 时的效果如下:

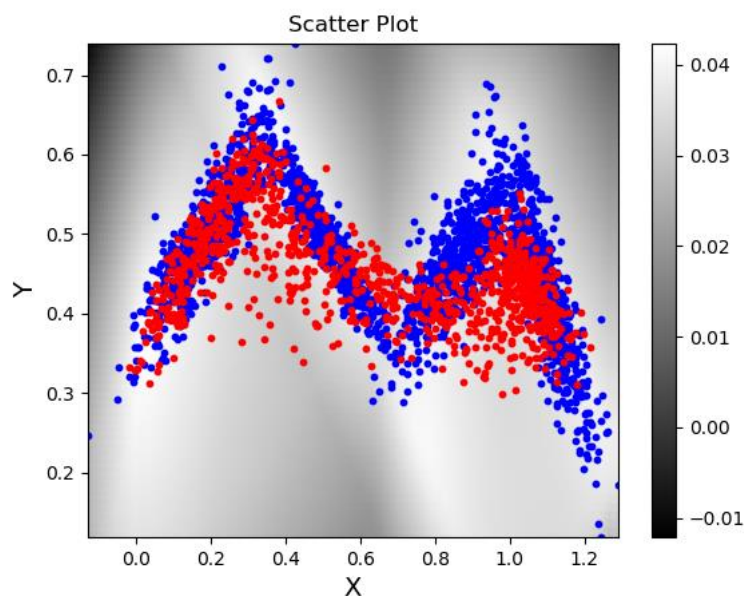


WGAN 虽然初始变化的很慢, 但最后生成的效果很好, 判断坐标点的准确率比 GAN 要高

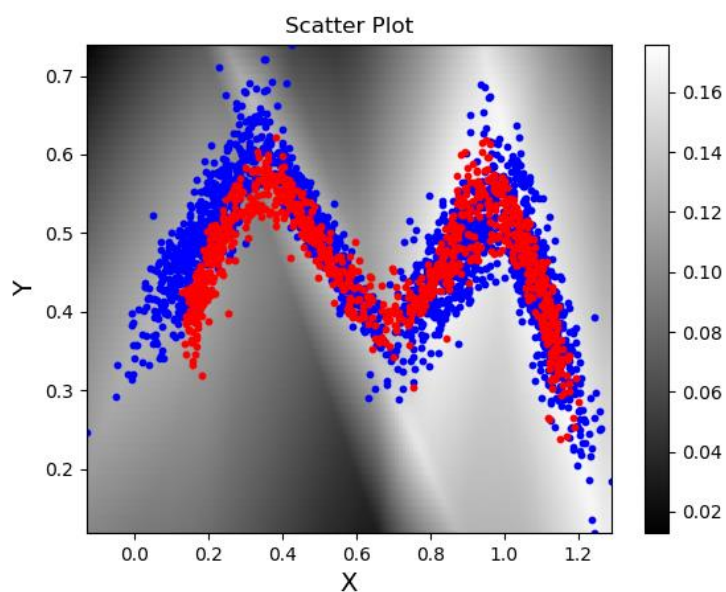
WGAN-GP:

WGAN-GP 只在 WGAN 基础上进行了更改，增加了一个惩罚项。惩罚项的计算在代码中体现了。下面看实验效果

在第 25 个 epoch 之后，生成的点就具有 M 的初始形状了：



最后，在第 100 个 epoch 后，已经能够很好地拟合了



总结：

三种 GAN 的效果相差不是很明显，最终都能成功拟合；WGAN 由于参数限制，拟合的速度较慢，但最终的效果很好。GAN 和 WGAN-GP 最终判断坐标点的准确率相比于 WGAN 低一些，但都不高；初步估计是训练的 epoch 比较小的原因。

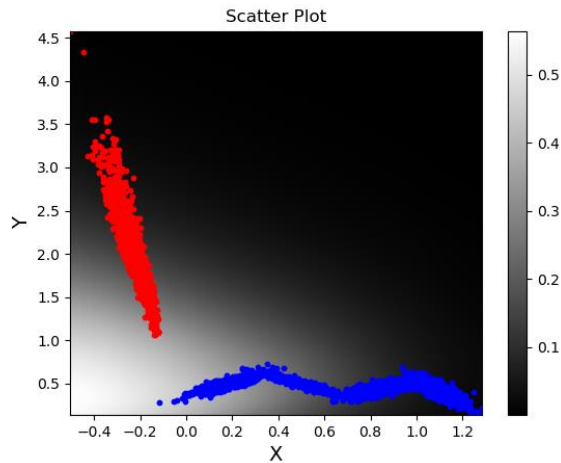
稳定性上，一旦拟合之后，三个网络的效果都不会变得很差，一直保持拟合状态。

对比不同优化器的影响

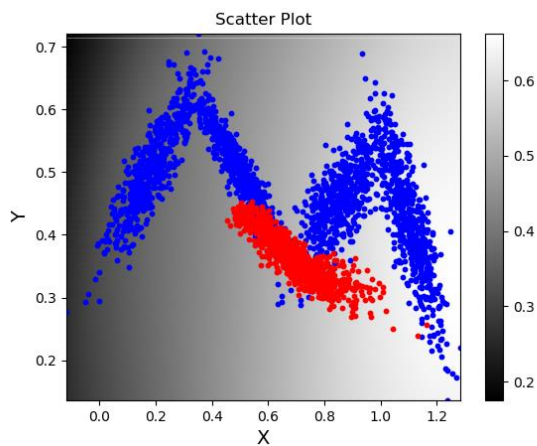
下面以 GAN 为例，对比 Adam、SGD、RMSprop 三个优化器

Adam:

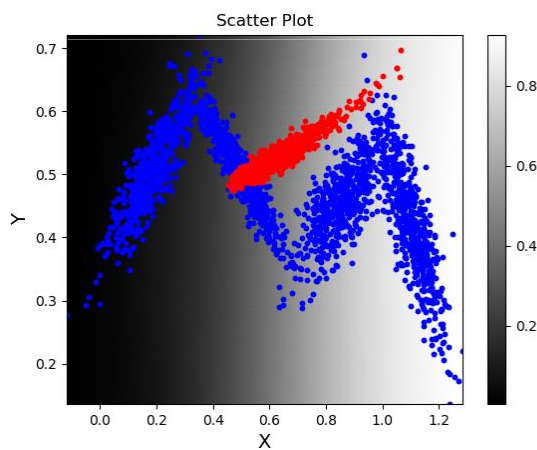
Adam 优化器的效果不好，很难拟合出来这个分布，初始时，训练 5 次后的结果：



训练 30 次时:

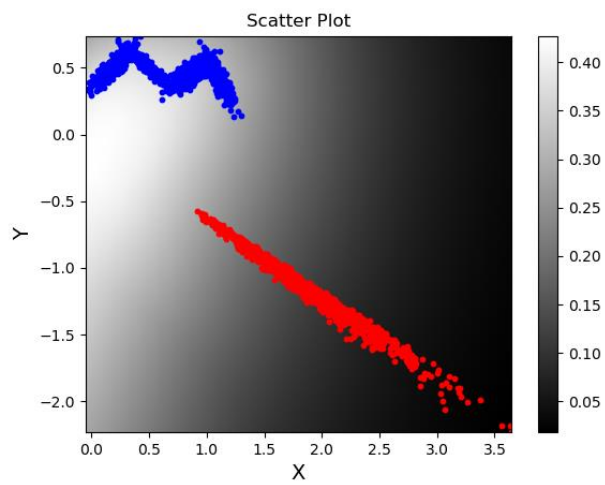


后面基本上是一“坨”的，生成的点分布，没有拐弯的地方，100 个 epoch 后的效果如下：

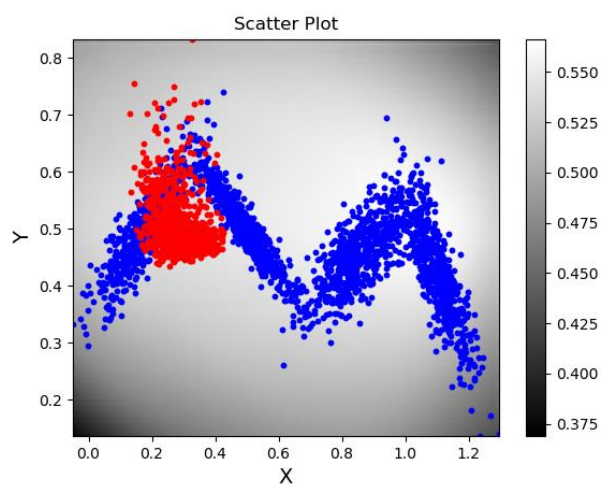


SGD:

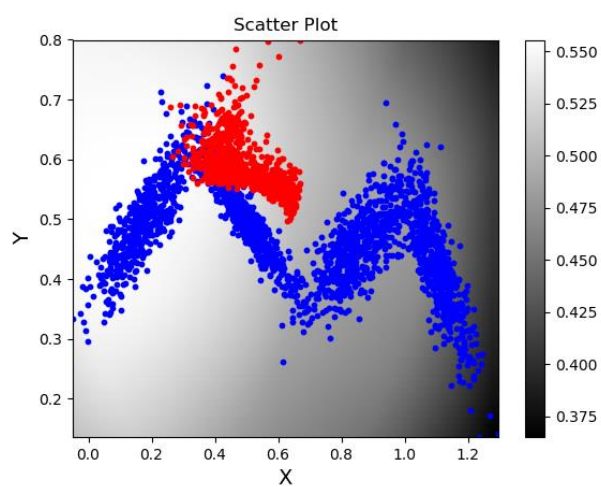
SGD 的效果也不好, 与 Adam 一样, 并不能成功拟合, 初始时, 训练 5 个 epoch 的效果:



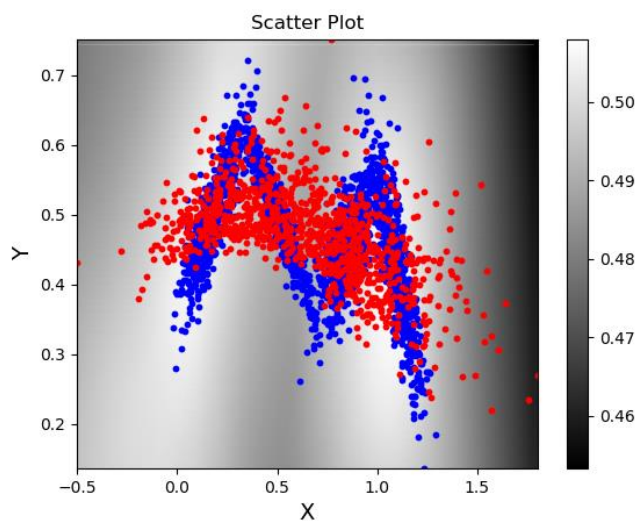
在第 30 个 epoch 时, 效果如下:



最后, 100 个 epoch 跑出的结果如下:



由于是怀疑 epoch 太小的原因, 我将 epoch 增大到了 1000, 效果如下:

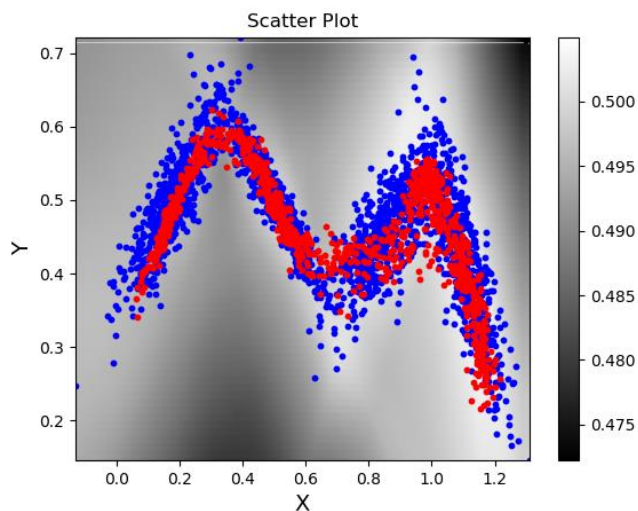


可以看到，虽然在不断地拟合，但点的分布并没有明显的拐点。

RMSprop:

RMSprop 的效果最好，能够成功拟合，而且速度很快。

在第 20 个 epoch 就已经能够拟合了，效果见上图。第 60 个 epoch 已经能够很好地拟合了：



总而言之，Adam 和 SGD 并不能成功拟合这次给定的分布，起码在有限的训练次数 (1000) 中不能成功拟合；而 RMSprop 优化器在第 60 次 epoch 就可以很好地拟合了。