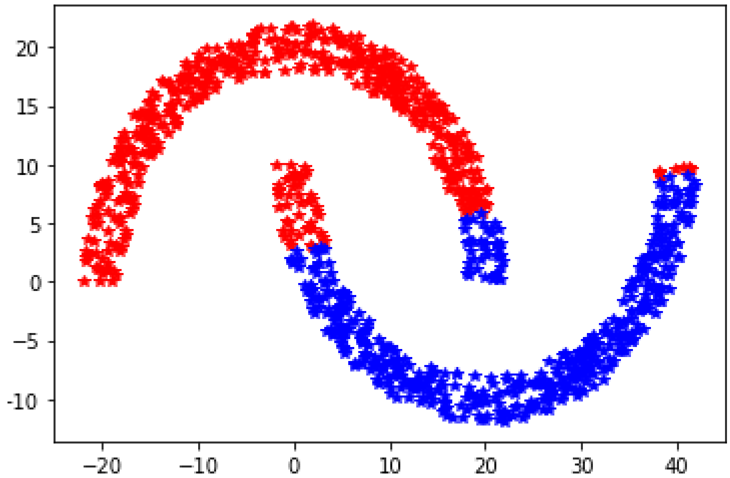
神经网络 作业4

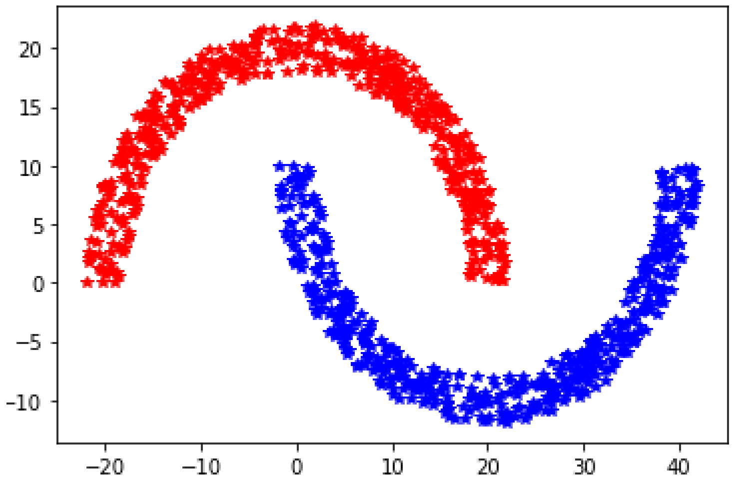
丁豪 人工智能学院

181220010@smail.nju.edu.cn

第一题

- 实验文件为 `第一题.ipynb` 。通过可视化的方式展示了单感知器、BP网络对双月数据集的分类情况。





- 分析：单层感知器的拟合能力相当于线性分类器，无法对线性不可分的双月数据集进行分类。而BP神经网络加入了隐藏层神经元，其中的激活函数采用了 `relu函数` 来去线性化，使网络的拟合能力超过了线性函数的限制，因此可以对非线性数据集进行拟合。

第二题

- 实验文件为 `第二题.ipynb` 。比较了不同隐层神经元数量的拟合正确率，之后又尝试了增加了隐藏层数量并采用交叉验证。
- 隐层神经元数量影响：隐藏层神经元数量越多，拟合能力越强，在充分训练的情况下精度越高。但是精度随神经元数量的提升速度呈下降趋势，这是因为已经能非常好的拟合训练集的时候额外提升拟合能力容易导致过拟合。

```
hidden: 1   acc: 0.304396
hidden: 5   acc: 0.805231
hidden: 10  acc: 0.930440
hidden: 20  acc: 0.954368
hidden: 50  acc: 0.954925
```

- 进一步提高方法：增加隐层神经元数量，增加层数，增加训练轮数，引入正则化减小过拟合。

```
0.9556135770234987
0.9869451697127938
0.9843342036553525
0.9842931937172775
0.9659685863874345
0.9869109947643979
0.9790575916230366
0.9764397905759162
0.9869109947643979
0.9764397905759162
hidden: 20,100,40 with 10_Fold_CrossValidation  acc: 0.959377
```

第三题

- 文件位于 `第三题.ipynb` 。采用pytorch实现，隐层激活函数用relu，训练器用adam，损失用mse，最终精度如下：

```
epoch 91, loss 0.0087
epoch 92, loss 0.0087
epoch 93, loss 0.0087
epoch 94, loss 0.0087
epoch 95, loss 0.0087
epoch 96, loss 0.0087
epoch 97, loss 0.0087
epoch 98, loss 0.0087
epoch 99, loss 0.0087
epoch 100, loss 0.0086
```

- 输入-隐藏，隐藏-输出层权值为

```
linear1.weight Parameter containing:
tensor([[-0.0263],
        [-0.6343],
        [ 0.4072],
        [-1.1352],
        [ 0.1665]], requires_grad=True)
linear1.bias Parameter containing:
tensor([-0.4353, -0.0133,  0.4695,  0.0069,  0.1663], requires_grad=True)
linear2.weight Parameter containing:
tensor([[ 0.0360,  0.0840,  0.4442, -0.2487,  0.4166]], requires_grad=True)
linear2.bias Parameter containing:
tensor([-0.1576], requires_grad=True)
```

- 因为采用relu函数，因此可以发现每一个神经元都存在一个由0转非0的分界点，具体数值如下（十分庞大）。最终的输出即这些神经元（分段函数）的组合，针对不同的输入，只有部分神经元有有效输出参与最终计算。

```
tensor([[0.0000e+00, 1.9793e+00, 0.0000e+00, 3.5733e+00, 0.0000e+00],
        [0.0000e+00, 1.9391e+00, 0.0000e+00, 3.5013e+00, 0.0000e+00],
        [0.0000e+00, 1.8988e+00, 0.0000e+00, 3.4292e+00, 0.0000e+00],
        [0.0000e+00, 1.8585e+00, 0.0000e+00, 3.3572e+00, 0.0000e+00],
        [0.0000e+00, 1.8183e+00, 0.0000e+00, 3.2851e+00, 0.0000e+00],
        [0.0000e+00, 1.7780e+00, 0.0000e+00, 3.2131e+00, 0.0000e+00],
        [0.0000e+00, 1.7378e+00, 0.0000e+00, 3.1410e+00, 0.0000e+00],
        [0.0000e+00, 1.6975e+00, 0.0000e+00, 3.0690e+00, 0.0000e+00],
        [0.0000e+00, 1.6573e+00, 0.0000e+00, 2.9969e+00, 0.0000e+00],
        [0.0000e+00, 1.6170e+00, 0.0000e+00, 2.9249e+00, 0.0000e+00],
        [0.0000e+00, 1.5768e+00, 0.0000e+00, 2.8528e+00, 0.0000e+00],
        [0.0000e+00, 1.5365e+00, 0.0000e+00, 2.7808e+00, 0.0000e+00],
        [0.0000e+00, 1.4963e+00, 0.0000e+00, 2.7087e+00, 0.0000e+00],
        [0.0000e+00, 1.4560e+00, 0.0000e+00, 2.6367e+00, 0.0000e+00],
        [0.0000e+00, 1.4158e+00, 0.0000e+00, 2.5646e+00, 0.0000e+00],
        [0.0000e+00, 1.3755e+00, 0.0000e+00, 2.4926e+00, 0.0000e+00],
        [0.0000e+00, 1.3352e+00, 0.0000e+00, 2.4205e+00, 0.0000e+00],
        [0.0000e+00, 1.2950e+00, 0.0000e+00, 2.3485e+00, 0.0000e+00],
        [0.0000e+00, 1.2547e+00, 0.0000e+00, 2.2764e+00, 0.0000e+00],
        [0.0000e+00, 1.2145e+00, 0.0000e+00, 2.2044e+00, 0.0000e+00],
        [0.0000e+00, 1.1742e+00, 0.0000e+00, 2.1323e+00, 0.0000e+00],
        [0.0000e+00, 1.1340e+00, 0.0000e+00, 2.0603e+00, 0.0000e+00],
        [0.0000e+00, 1.0937e+00, 0.0000e+00, 1.9882e+00, 0.0000e+00],
        [0.0000e+00, 1.0535e+00, 0.0000e+00, 1.9162e+00, 0.0000e+00],
        [0.0000e+00, 1.0132e+00, 0.0000e+00, 1.8441e+00, 0.0000e+00],
        [0.0000e+00, 9.7296e-01, 0.0000e+00, 1.7721e+00, 0.0000e+00],
        [0.0000e+00, 9.3271e-01, 0.0000e+00, 1.7000e+00, 0.0000e+00],
        [0.0000e+00, 8.9245e-01, 0.0000e+00, 1.6280e+00, 0.0000e+00],
        [0.0000e+00, 8.5220e-01, 0.0000e+00, 1.5559e+00, 0.0000e+00],
        [0.0000e+00, 8.1194e-01, 0.0000e+00, 1.4839e+00, 0.0000e+00],
        [0.0000e+00, 7.7169e-01, 0.0000e+00, 1.4118e+00, 0.0000e+00],
        [0.0000e+00, 7.3144e-01, 0.0000e+00, 1.3398e+00, 0.0000e+00],
        [0.0000e+00, 6.9118e-01, 1.7208e-02, 1.2677e+00, 0.0000e+00],
        [0.0000e+00, 6.5093e-01, 4.3053e-02, 1.1957e+00, 0.0000e+00],
        [0.0000e+00, 6.1067e-01, 6.8898e-02, 1.1237e+00, 2.5408e-03],
```

```
        [0.0000e+00, 5.7042e-01, 9.4743e-02, 1.0516e+00, 1.3106e-02],
        [0.0000e+00, 5.3017e-01, 1.2059e-01, 9.7955e-01, 2.3671e-02],
        [0.0000e+00, 4.8991e-01, 1.4643e-01, 9.0751e-01, 3.4236e-02],
        [0.0000e+00, 4.4966e-01, 1.7228e-01, 8.3546e-01, 4.4802e-02],
        [0.0000e+00, 4.0940e-01, 1.9812e-01, 7.6341e-01, 5.5367e-02],
        [0.0000e+00, 3.6915e-01, 2.2397e-01, 6.9136e-01, 6.5932e-02],
        [0.0000e+00, 3.2890e-01, 2.4981e-01, 6.1931e-01, 7.6497e-02],
        [0.0000e+00, 2.8864e-01, 2.7566e-01, 5.4726e-01, 8.7063e-02],
        [0.0000e+00, 2.4839e-01, 3.0150e-01, 4.7521e-01, 9.7628e-02],
        [0.0000e+00, 2.0813e-01, 3.2735e-01, 4.0317e-01, 1.0819e-01],
        [0.0000e+00, 1.6788e-01, 3.5319e-01, 3.3112e-01, 1.1876e-01],
        [0.0000e+00, 1.2763e-01, 3.7904e-01, 2.5907e-01, 1.2932e-01],
        [0.0000e+00, 8.7372e-02, 4.0488e-01, 1.8702e-01, 1.3989e-01],
        [0.0000e+00, 4.7118e-02, 4.3073e-01, 1.1497e-01, 1.5045e-01],
        [0.0000e+00, 6.8638e-03, 4.5657e-01, 4.2923e-02, 1.6102e-01],
        [0.0000e+00, 0.0000e+00, 4.8242e-01, 0.0000e+00, 1.7158e-01],
        [0.0000e+00, 0.0000e+00, 5.0826e-01, 0.0000e+00, 1.8215e-01],
        [0.0000e+00, 0.0000e+00, 5.3411e-01, 0.0000e+00, 1.9271e-01],
        [0.0000e+00, 0.0000e+00, 5.5995e-01, 0.0000e+00, 2.0328e-01],
        [0.0000e+00, 0.0000e+00, 5.8580e-01, 0.0000e+00, 2.1385e-01],
        [0.0000e+00, 0.0000e+00, 6.1165e-01, 0.0000e+00, 2.2441e-01],
        [0.0000e+00, 0.0000e+00, 6.3749e-01, 0.0000e+00, 2.3498e-01],
        [0.0000e+00, 0.0000e+00, 6.6334e-01, 0.0000e+00, 2.4554e-01],
        [0.0000e+00, 0.0000e+00, 6.8918e-01, 0.0000e+00, 2.5611e-01],
        [0.0000e+00, 0.0000e+00, 7.1503e-01, 0.0000e+00, 2.6667e-01],
        [0.0000e+00, 0.0000e+00, 7.4087e-01, 0.0000e+00, 2.7724e-01],
        [0.0000e+00, 0.0000e+00, 7.6672e-01, 0.0000e+00, 2.8780e-01],
        [0.0000e+00, 0.0000e+00, 7.9256e-01, 0.0000e+00, 2.9837e-01],
        [0.0000e+00, 0.0000e+00, 8.1841e-01, 0.0000e+00, 3.0893e-01],
        [0.0000e+00, 0.0000e+00, 8.4425e-01, 0.0000e+00, 3.1950e-01],
        [0.0000e+00, 0.0000e+00, 8.7010e-01, 0.0000e+00, 3.3006e-01],
        [0.0000e+00, 0.0000e+00, 8.9594e-01, 0.0000e+00, 3.4063e-01],
        [0.0000e+00, 0.0000e+00, 9.2179e-01, 0.0000e+00, 3.5119e-01],
        [0.0000e+00, 0.0000e+00, 9.4763e-01, 0.0000e+00, 3.6176e-01],
        [0.0000e+00, 0.0000e+00, 9.7348e-01, 0.0000e+00, 3.7232e-01],
        [0.0000e+00, 0.0000e+00, 9.9932e-01, 0.0000e+00, 3.8289e-01],
        [0.0000e+00, 0.0000e+00, 1.0252e+00, 0.0000e+00, 3.9345e-01],
        [0.0000e+00, 0.0000e+00, 1.0510e+00, 0.0000e+00, 4.0402e-01],
        [0.0000e+00, 0.0000e+00, 1.0769e+00, 0.0000e+00, 4.1458e-01],
        [0.0000e+00, 0.0000e+00, 1.1027e+00, 0.0000e+00, 4.2515e-01],
        [0.0000e+00, 0.0000e+00, 1.1285e+00, 0.0000e+00, 4.3572e-01],
        [0.0000e+00, 0.0000e+00, 1.1544e+00, 0.0000e+00, 4.4628e-01],
        [0.0000e+00, 0.0000e+00, 1.1802e+00, 0.0000e+00, 4.5685e-01],
        [0.0000e+00, 0.0000e+00, 1.2061e+00, 0.0000e+00, 4.6741e-01],
        [0.0000e+00, 0.0000e+00, 1.2319e+00, 0.0000e+00, 4.7798e-01],
        [0.0000e+00, 0.0000e+00, 1.2578e+00, 0.0000e+00, 4.8854e-01],
        [0.0000e+00, 0.0000e+00, 1.2836e+00, 0.0000e+00, 4.9911e-01],
        [0.0000e+00, 0.0000e+00, 1.3095e+00, 0.0000e+00, 5.0967e-01],
        [0.0000e+00, 0.0000e+00, 1.3353e+00, 0.0000e+00, 5.2024e-01],
        [0.0000e+00, 0.0000e+00, 1.3612e+00, 0.0000e+00, 5.3080e-01],
        [0.0000e+00, 0.0000e+00, 1.3870e+00, 0.0000e+00, 5.4137e-01],
        [0.0000e+00, 0.0000e+00, 1.4128e+00, 0.0000e+00, 5.5193e-01],
```

```
        [0.0000e+00, 0.0000e+00, 1.4387e+00, 0.0000e+00, 5.6250e-01],
        [0.0000e+00, 0.0000e+00, 1.4645e+00, 0.0000e+00, 5.7306e-01],
        [0.0000e+00, 0.0000e+00, 1.4904e+00, 0.0000e+00, 5.8363e-01],
        [0.0000e+00, 0.0000e+00, 1.5162e+00, 0.0000e+00, 5.9419e-01],
        [0.0000e+00, 0.0000e+00, 1.5421e+00, 0.0000e+00, 6.0476e-01],
        [0.0000e+00, 0.0000e+00, 1.5679e+00, 0.0000e+00, 6.1532e-01],
        [0.0000e+00, 0.0000e+00, 1.5938e+00, 0.0000e+00, 6.2589e-01],
        [0.0000e+00, 0.0000e+00, 1.6196e+00, 0.0000e+00, 6.3645e-01],
        [0.0000e+00, 0.0000e+00, 1.6454e+00, 0.0000e+00, 6.4702e-01],
        [0.0000e+00, 0.0000e+00, 1.6713e+00, 0.0000e+00, 6.5758e-01],
        [0.0000e+00, 0.0000e+00, 1.6971e+00, 0.0000e+00, 6.6815e-01],
        [0.0000e+00, 0.0000e+00, 1.7230e+00, 0.0000e+00, 6.7872e-01],
        [0.0000e+00, 0.0000e+00, 1.7488e+00, 0.0000e+00, 6.8928e-01]],
```