

“深度学习和计算机视觉（实践）”课程实验三： 单隐藏层神经网络的训练

廖振宇

日期：2021 年 9 月 24 日

1 实验背景

课程内容回顾 在理论课程中，我们讨论了神经网络的基本结构，具体讨论了关于激活函数、神经网络架构、基于正向和反向传播的网络训练、不同的正则化方式等一系列神经网络训练的问题和对应策略。

实验内容 不使用 Keras, Tensorflow 或 Pytorch 等框架，仅使用 Numpy, Scipy 和 Matplotlib 等 Python 常用科学计算库，完成单隐藏层的全连接神经网络（和之后要讲的卷积神经网络形成对比），实现 Scikit-learn 里的 half moons 数据集的分类。

2 实验要求

- 使用 Scikit-learn 加载 half moon 数据，实现数据的可视化
- 构建单隐藏层的全连接神经网络，参考课件的内容，推导正、反向传播并给出代码实现
- 考虑不同的激活函数、不同的隐藏层宽度对网络训练和性能的影响
- 需要具备的功能或者模块：数据读取、加载，正反向传播，使用梯度下降训练模型（可采用 mini-batch 的随机梯度下降方法）¹，模型训练、测试性能指标的显示和评估
- PPT 汇报（每组 3min），提交 2-4 页实验报告，需简要叙述方法原理、实验步骤、方法参数讨论、实验结果；需明确说明组员分工、给出组内排名（可标注同等贡献 #）。

3 参考资料

- Scikit-learn 的 half moons 数据集：https://scikit-learn.org/stable/modules/generated/sklearn.datasets.make_moons.html

¹模型参数需要同时一起更新，可以尝试不同的停止训练的策略，注意随机初始化的选择。