

深度学习导论 DS2001.01.2024SP——实验一

实验要求

使用 pytorch 或者 tensorflow 手写一个前馈神经网络，用于近似以下函数：

$$y = \log_2(x) + \cos\left(\frac{\pi x}{2}\right), x \in [1,16]$$

研究数据量、网络深度、学习率、网络宽度、激活函数对模型性能的影响。

实验步骤

1. **网络框架**：要求选择 pytorch 或 tensorflow 其中之一，依据官方网站的指引安装包。若你需要使用 GPU，可能还需安装 CUDA 驱动。本次实验仅利用 CPU 也可以完成，但仍强烈推荐大家安装 GPU 版本，以满足后续实验需求。

2. **数据生成**：本次实验的数据集仅需使用程序自动生成，即在取值范围内随机采样作为 x 值，并计算相应的 y 值。为统一规范，本次实验要求在 $x \in [1,16]$ 范围内均匀采样获得 N 个样本的数据集，再按**训练集：验证集：测试集=8: 1: 1**的比例随机划分成三个**互不相交**的部分，并且在调参分析中**固定**三个部分的组成。

要求：请在 $N=200/2000/10000$ 三种不同数据量情况下开展实验。

3. **模型搭建**：采用 pytorch 或 tensorflow 所封装的 module 编写模型，例如 `torch.nn.Linear()`, `torch.nn.ReLU()`等，无需手动完成底层 forward、backward 过程。

4. **模型训练**：将生成的**训练集**输入搭建好的模型进行前向的 loss 计算和反向的梯度传播，从而训练模型，同时也建议使用网络框架封装的 optimizer 完成参数更新过程。

提示：模型搭建和训练环节在 PPT 上有示例可供参考。

5. **调参分析**：将训练好的模型在**验证集**上进行测试，以均方误差(MSE)作为网络性能指标。然后，对**网络深度、学习率、网络宽度、激活函数**等模型的超参数进行调整后，再重新训练、测试，并分析对模型性能的影响。

要求：为便于直观比较，请将验证集原始样本点和预测点利用 matplotlib 绘图进行可视化。

6. **测试性能**：选择你认为最合适的（例如，在验证集上表现最好的）一组超参数，重新训练模型，并在**测试集**上测试（**注意，此处应是你的实验中唯一一次在测试集上的测试**），并记录测试的结果（MSE）。

实验提交

实验一截止时间：**3月31日 23:59:59**，线下完成代码检查（关键代码讲解+运行展示+结果展示），并需在 bb 系统提交源代码及实验报告，具体要求如下：

- 全部文件打包在一个压缩包内，压缩包命名为：**学号-姓名-ex1.zip**
- 代码仅包含.py 文件，请勿包含实验中间结果（例如中间保存的数据集等）；如果有多个代码文件，放在 **src/**文件夹内
- 实验报告提交为.pdf 格式，包含学号、姓名，内容包括简要的实验过程、关键代码展示，重点关注**对超参数的实验分析以及测试集上的实验结果**。