

Assignment2 评分准则

Part1: 30 分

Task1&2 (20 分)

实现 pytorch mlp 代码 **2 分**；若代码没实现完全，则 task1&2 的代码及报告部分**酌情给分**。

请在 ipynb 嵌入 numpy mlp (assignment1) 代码并与 pytorch mlp 比较：

- pytorch mlp 在 make moons 数据集上能够正确训练且测试，并且输出合理结果 **2 分**，accuracy 过低则酌情给分；若 pytorch mlp 不能运行，则整个 Task1&2 **1 分** 辛苦分，若没有代码或代码极少，则 **0 分**。
- **再次**在 make moons 数据集上运行 assignment1 的 numpy mlp 的代码，并给出结果。若结果与 pytorch mlp 的结果差距过大，则扣分。
- 请在额外两个数据集上比较 pytorch mlp 与 numpy mlp，完整比较一个数据集 **3 分**。同上，结果差距过大则**扣分**，
- 在以上对两个 mlp 在三个数据集上的比较实验中，超参数**均须设置一致**，超参数不同则该部分报告**不给分**，代码酌情给分。

报告 **10 分**：

- 详细报告包含上述及作业文档的所有要求，需要包含所有步骤，包括实现方法及思路、超参数设置**等**，详情请参考作业文档，**不详细视完整程度扣分**。
- 所有的运行过程都需要 accuracy 曲线图片和 loss 曲线图片：少一条曲线报告**扣 1 分**。该部分报告分扣完为止。
- 请对所有的结果进行解释和分析，只有图片没解释**不给分**，没有图片**不给分**，根据分析的质量给分。

Task3

实现代码且可以在 ipynb 正确运行 (**4 分**)，虽然 mlp 对 cifar10 分类效果**不是特别好**，但是我们对效果会有个预期，如果结果过于离谱会认定为代码实现有问题，这个 task **酌情给分**。

报告 **6 分**：

- 包含 task3 所有内容，需要包括实现方法、超参数设置等，并附上相关曲线图，图片的要求参照上文。
- 需要对最终结果进行详细分析，根据分析的质量给分。

Part2: 30 分

实现 CNN 代码且能正确的在 ipynb 上完整运行 4 分。若 CNN 不能运行，则 Part2 1 分

除了默认的优化器与学习率，需要尝试另外两个优化器进行训练 一个额外优化器 2 分（最高 4 分），；在使用默认优化器的条件下尝试使用不同的学习率进行训练，一个学习率 2 分（最高 4 分），并比较结果、给出曲线图、对最终结果解释。

报告 18 分：

- 解释如何实现这部分，包含详细的过程，需要有每个 layer 的通道数，输入输出大小的计算等；
- 需要结果的曲线图，图片要求参照 Part1。
- 需要对最终结果进行详细分析，根据分析的质量给分。但如果报告只放图片没有解释或只有解释没有图片不给分。

Part3: 40 分

实现 rnn 代码并可以在 ipynb 上正确完整运行，默认设置长度为 5，（15 分）。因为默认长度为 5，所以效果会很好，如果效果没达到预期则视为该部分实现不完整，整个 part 酌情给分。作业文档中明确表示不能使用 torch.nn.RNN 或 torch.nn.LSTM，若使用了，则不得分。

除了长度 5，至少需要额外三个不低于 5 的测试结果，6 分，少一个扣 2 分，长度为 10，15，20。

报告 18 分：

- 详细包含上述及作业文档的所有要求，包括训练过程，写代码的过程，要求有公式解释等。
- 要求结果曲线图，要求参考 Part1。
- 要求对不同长度结果进行解释分析，只有图片没解释不给分，没有图片不给分，根据分析的质量给分。

补充说明

所有曲线图须包括 loss 图 与 accuracy 图， 少了则该部分扣分。

part1、2 所有曲线图均需包含测试集与训练集，可以不在同一张图片但一定要有，若没有测试集或训练集则**该部分扣分**；如果训练集包含测试集或验证集，**该部分扣分**。

如果只有代码但报告没有相应的结果描述和分析，该部分代码**酌情给分**，报告相应部分不给分。

助教会随机抽取作业进行结果复现，若复现结果与报告结果差距过大，则酌情给分，允许线下 argue。

此准则为对作业文档的补充说明，须同时满足作业文档与此准则。同时报告**须**为单独的 pdf 文件。