

HW2 Report

卢华昊 2017011477 计 72

一、 实验目标

- a) 使用 MLP、CNN 等网络进行分类。
- b) 观察实验结果并且回答思考题。

二、 网络结构及参数设置

1. MLP 使用双线性层的结构，第一层为 512 个神经元，第二层为 10 个神经元。激活函数为 Relu。

Lr=3e-4, Epoch=100, dropout_rate=0.4

2. CNN 使用双层卷积与最大池化的结构。第一层卷积 filters 数量为 32, kernal_size 为 8, 最大池化池化尺寸为 4x4, 步长为 1; 第二层卷积与第一层卷积相同, 最大池化与第一层相同。

Lr=3e-4, Epoch=100, dropout_rate=0.2

三、 实验结果

1. 在 MLP 结构中, 在 37 个 Epoch 之后达到最优效果, 测试集上准确率为 55.84%
2. 在 CNN 结构中, 在 31 个 Epoch 之后达到最优, 测试集上准确率达到 69.07%

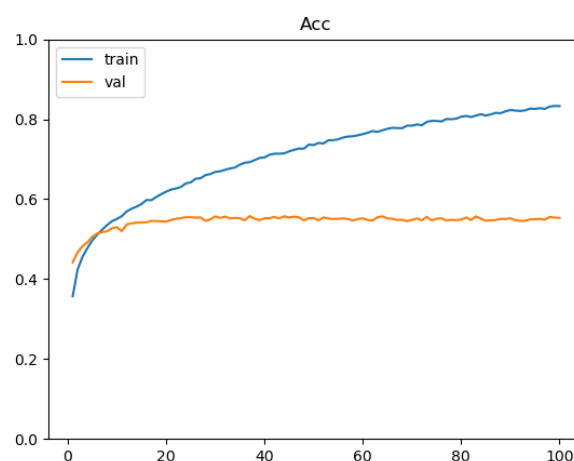
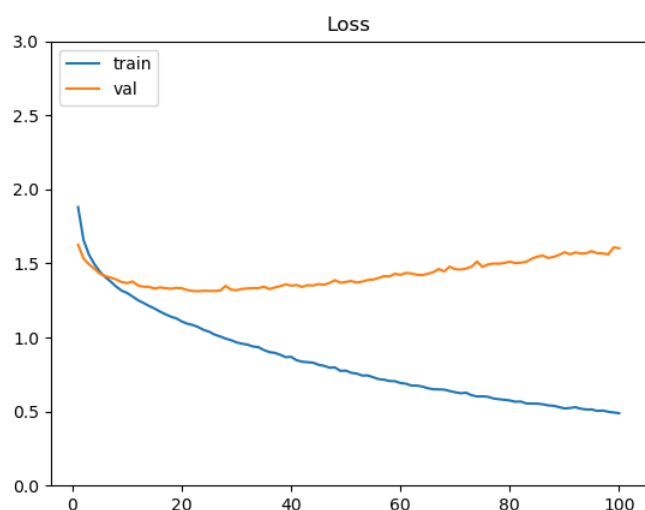
四、 问题回答及附录

1. 在训练时 `is_train=True`, 测试时 `is_train=False`, 这一选项涉及 `batch_normalisation`、`dropout` 等算法的使用。训练时 `reuse=False`, 测试时 `reuse=True`, 这一选项设计是否重用参数。

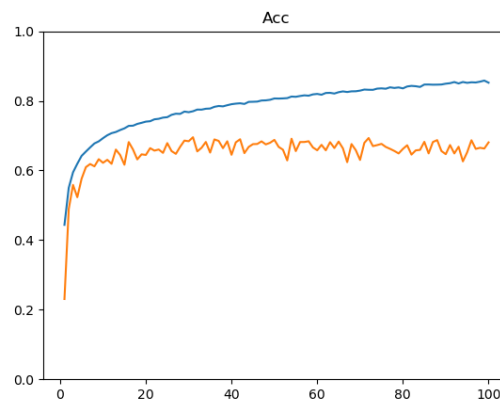
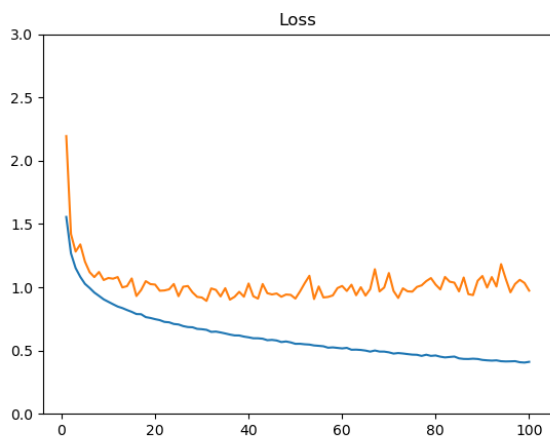
`Batch_norm` 在训练时和测试时使用不同的计算策略, 训练时使用参与当前训练的 mini-batch 进行计算, 而在测试时使用之前训练集所有计算结果的均值进行 `normalize`。

`Dropout` 在训练时随机扔掉一些神经元, 而在测试时则不会扔掉神经元, 而是使用全部的神经元参与计算。

2.



MLP 模型, $dropout=0.4$. $lr=3e-4$



CNN 模型, dropout=0.2, lr=3e-4

3. MLP 与 CNN 的不同点:

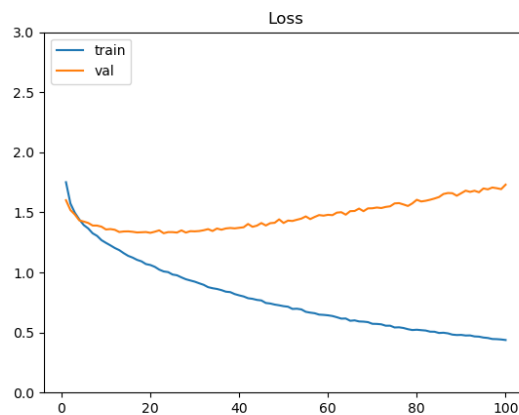
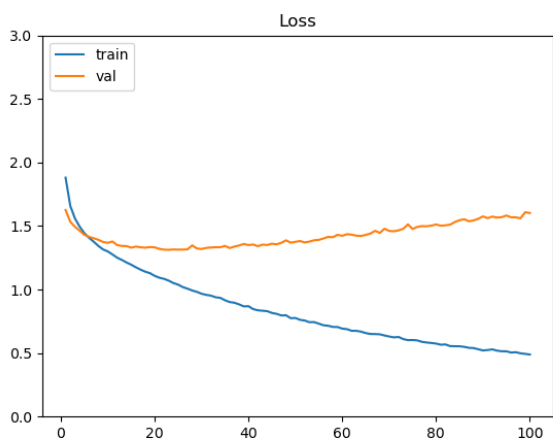
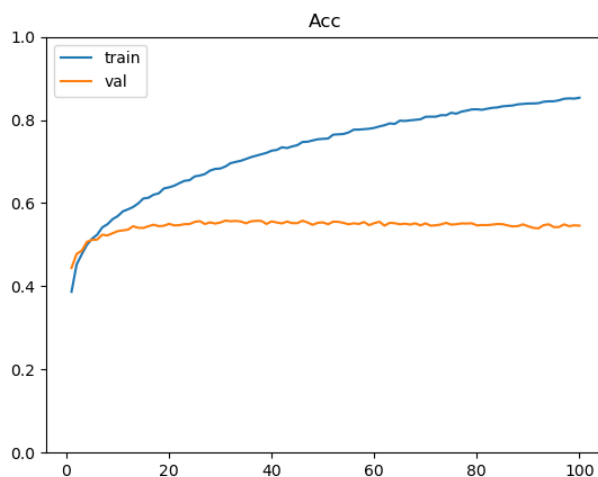
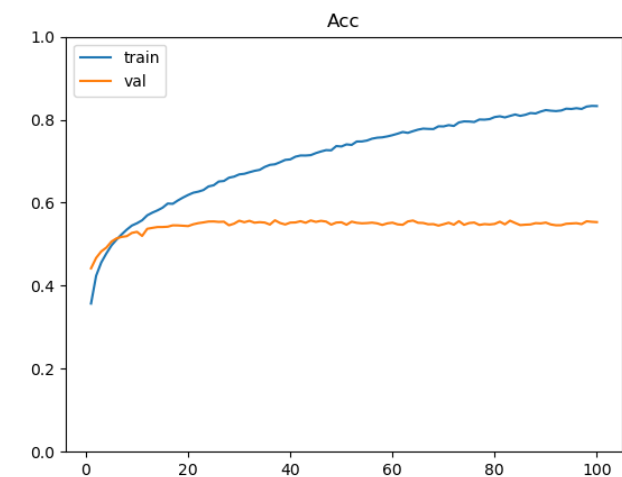
- a) CNN 在测试集上的准确率超过了 MLP。MLP 的准确率最高为 55%左右, 而 CNN 的准确率可以达到 68%左右。
- b) MLP 在训练集上对 Dropout Rate 的变化更敏感。从下文中的表格中可以看出, 同样 Dropout Rate 的变化下, MLP 在训练集上的结果变化更加明显。
- c) CNN 在测试集上对 Dropout Rate 的变化更敏感。从下文中的表格中可以看出, 同样 Dropout Rate 的变化下, CNN 在测试集上的结果变化更加明显。
- d) Batch Norm 有无对 CNN 影响更显著。从下文可以看出, Batch Norm 的有无显著影响了 CNN 在测试集上的稳定性, 而 MLP 则影响不明显。

4. Batch Normalization 的效果:

- a) MLP:

由于在进行 Batch Norm 时经过测试，dropout=0.4 时训练效果最好，于是选择在 dropout=0.4 的情况下进行测试。

经过测试发现，是否进行 Batch Normalization 对训练结果无显著影响，最终准确率均为 55% 左右；测试集上达到最高点的速度也近似。

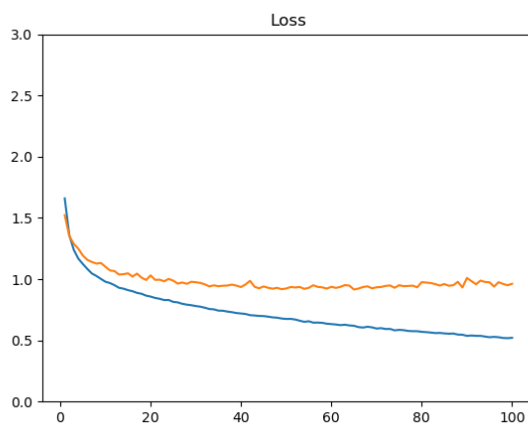
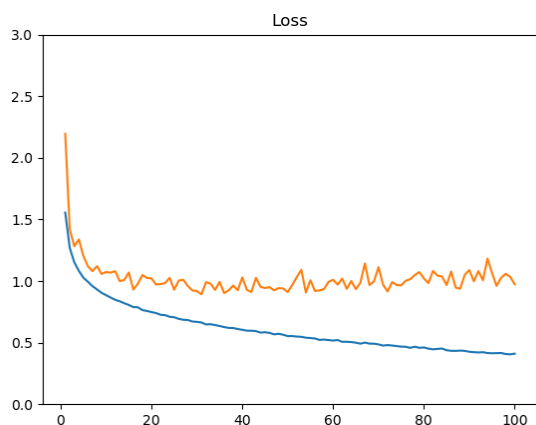
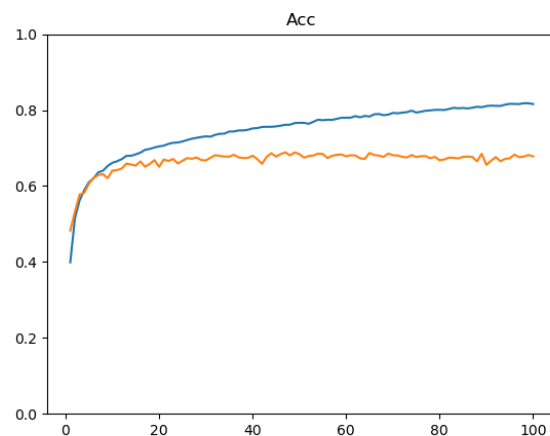
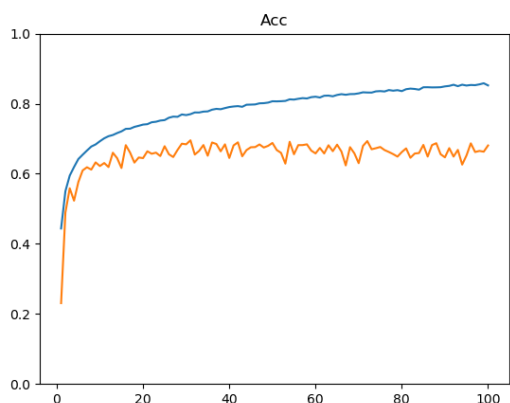


左图为 BatchNorm 时情况，右图为不进行 BatchNorm 时情况

b) CNN:

由于在进行 Batch Norm 时经过测试，dropout=0.2 时训练效果最好，于是选择在 dropout=0.2 的情况下进行测试。

依然没有发现明显的改变，最终准确率均为 69%左右。反而 Batch Norm 后在测试集上训练结果波动更大。

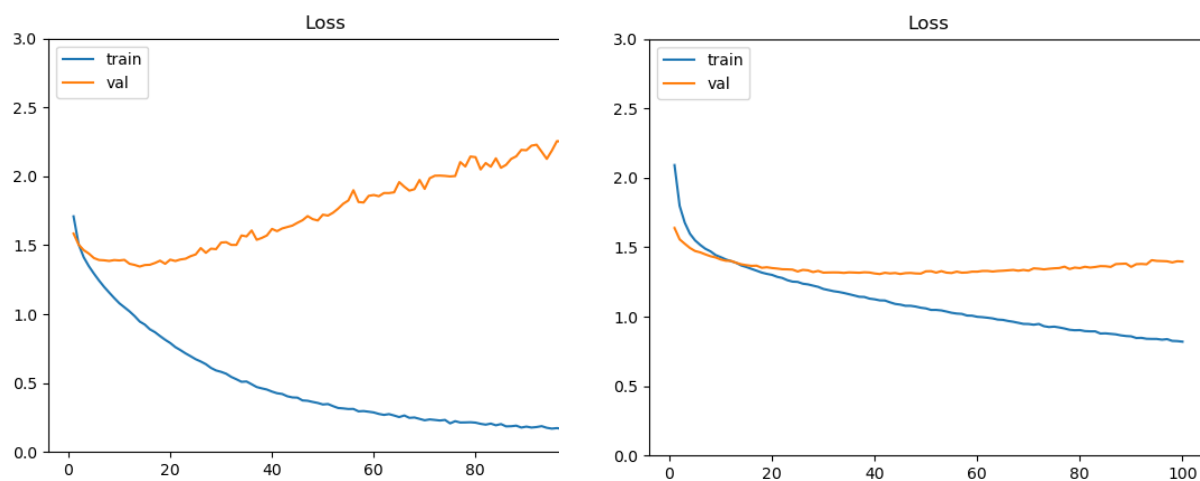


左图为 Batch Norm，右图不适用 Batch Norm

5. Dropout Rate 的大小影响在训练时扔掉神经元的概率，这一值越大，扔掉神经元的概率就越大。

在进行训练的过程中，我发现随着 dropout rate 的增加：

a) val_loss 在到达最低点后重新上升的速度变慢，同时上升的幅度更小，这一曲线更加平缓。



ML 左图为 rate=0.1 时的 loss 曲线，右图 rate=0.6

b) 训练集上的准确率和 loss 在同样的 Epoch 后也逐渐下降；

c) 在测试集上的准确率有小幅提升。

以下是测试情况的统计表格。

最高准确率(100 个 Epoch 之内)及达到验证集最佳所需 Epoch

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
MLP_train	94.67%	91.48%	87.66%	83.29%	77.74%	71.96%	64.44%	56.01%
MLP_val	54.93%	54.37%	55.22%	55.84%	55.55%	55.32%	54.99%	54.62%
MLP_epoch	19	43	37	37	67	64	77	92
CNN_train	88.07%	85.21%	83.16%	81.40%	79.26%	79.40%	74.00%	69.55%
CNN_val	67.08%	69.07%	68.58%	67.75%	66.41%	66.38%	54.75%	45.87%
CNN_epoch	33	31	50	67	42	95	54	87

6. 在训练的开始部分，训练集与测试集上的 loss 类似，相差不大；而在训练多轮 Epoch 后之后，训练集上的 loss 持续减小，而测试集上的 loss 先保持不变随后开始上升。

这可能是因为模型在训练集上出现了过拟合现象，因此我们对 dropout 的概率进行了调整与测试，这可能会减轻过拟合现象。具体的测试结果可以见上表。