

第 3 次作业报告

张家治 (1601110508)

本次测试的网络共有 4 个 ReLU 隐层，神经元数由前向后分别为 200、100、60、30，输出层为 softmax 层，使用 Adam 优化方法进行优化，学习率基于指数递减。仍基于 MNIST 数据集测试。

分别测试了加 Batch Normalization 和不加 Batch Normalization 情况下，0.3、0.03、0.003 的最大学习率的性能。结果显示，6 个网络准确率都能在 2 个 epoch 之内收敛到 0.95 以上。其中，加入 Batch Normalization 层的收敛速度快于所有不加 Batch Normalize 的收敛速度，且准确率也比其高 2 个百分点，而对于加入 Batch Normalization 层的网络，最大学习率 0.003 的收敛速度比 0.03 快，而 0.03 比 0.3 快，对于不加 Batch Normalization 的网络也是如此。

结论：Batch Normalization 能显著改善网络的准确率，并且能大大加速网络的收敛速度。

