

深度学习与神经网络第一次课程项目

王逸群 19307110397

2022.3.18

目标：对输入的 16×16 手写数字灰度图进行识别，输出 $y \in \{-1, 1\}^{10}$ ；

数据集：存储于 `digits.mat`；

激活函数：Tanh 函数；

损失函数：平方损失函数；

优化方法：随机梯度下降法。

模型评估：在测试集上的错误率为 47%，在验证集上的错误率如图 1。

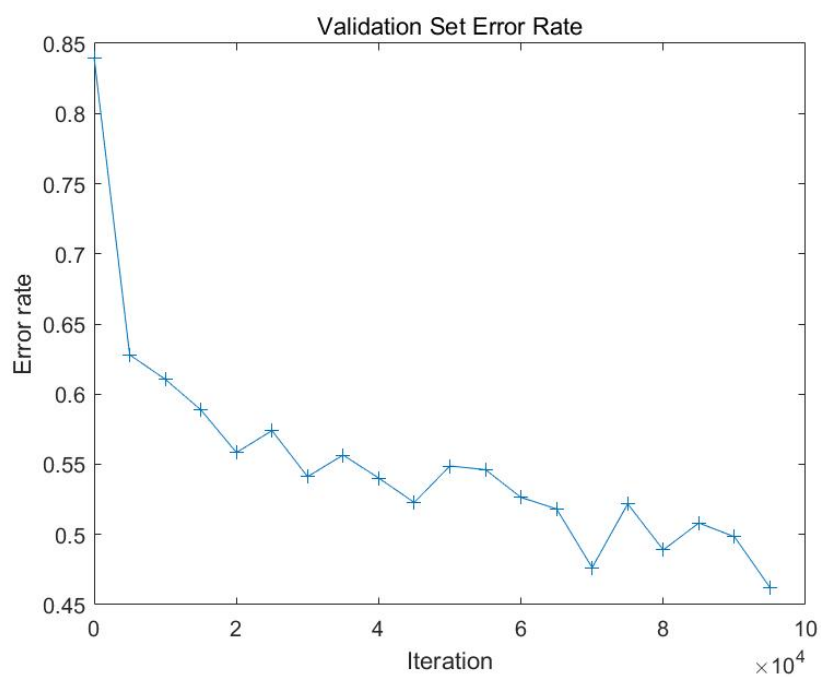


图 1: 原始模型在验证集上的错误率

1 网络结构

程序中，变量`nHidden`为一个向量，其维度表示网络隐藏层的个数，各个分量表示每个隐藏层的神经元个数。

基础模型中，`nHidden = 10`，将其修改得到的验证集和测试集错误率如图 2 和表 1 所示。

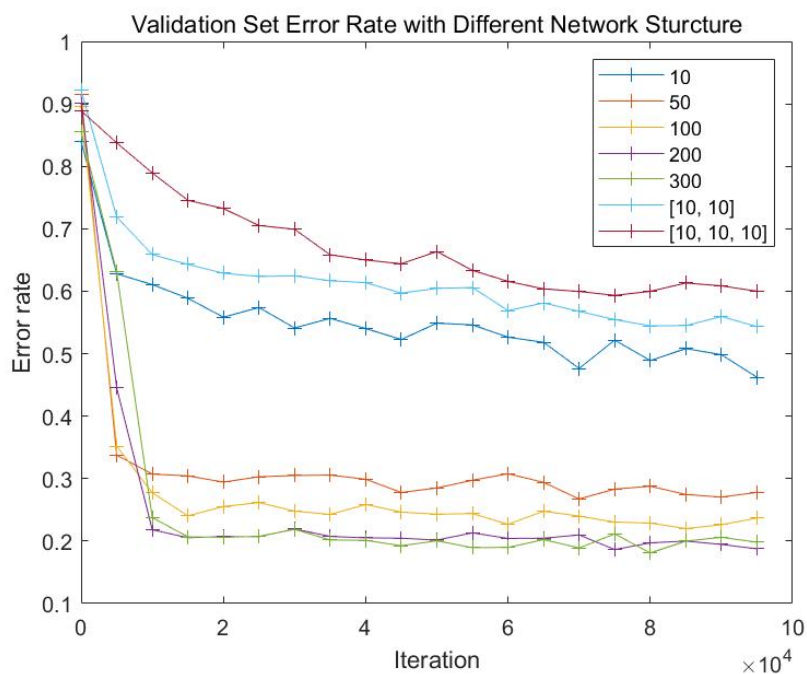


图 2: 不同网络结构的模型在验证集上的错误率

可以看到，随着单层网络神经元个数的增加，模型在测试集上的错误率逐渐减小，但是减小的程度趋缓。鉴于训练模型的时间成本随着网络神经元个数的增加而增加，神经元个数的选取是精度与成本的权衡。

另一方面，随着网络隐藏层个数的增加，模型在测试集上的错误率逐渐增加，可能的原因是出现了过拟合现象。

网络结构	测试集错误率
nHidden = 10;	0.470
nHidden = 50;	0.255
nHidden = 100;	0.228
nHidden = 200;	0.181
nHidden = 300;	0.177
nHidden = [10, 10];	0.536
nHidden = [10, 10, 10];	0.562

表 1: 不同网络结构的模型在测试集上的错误率