

1. LcdD 调整过程：

version 1：

学习率：0.03

隐藏层：18

次数：50000

均方误差：

0.0011142112208825993（训练集）

0.06301846724409145（验证集）

version 2：

学习率：0.06

隐藏层：18

次数：5000000

均方误差：

0.03219604135098666

0.051132129937386264

version 3：

学习率：0.08

隐藏层：18

次数：50000

均方误差：

0.04719928841996844

0.06657691145309054

version 4：

学习率：0.02

隐藏层：18

次数：50000

均方误差：

0.06692508063112781

0.07141475046000785

version 5：

学习率：0.05

隐藏层：12

次数：50000

均方误差：

0.04140914752528202

0.07764357385528765

version 6：

学习率：0.05

隐藏层：24

次数：50000

均方误差：

0.030971774214319087

0.05414234837094316

version 7 :

学习率 : 0.05

隐藏层 : 24

次数 : 100000

均方误差 :

0.010050468258397722

0.08911713140021688

version 8 :

学习率 : 0.05

隐藏层 : 24

次数 : 100000

均方误差 :

0.0014341084938368145

0.052819773328487814

version 9 :

先学习率为 0.05 训练 50000 次, 在学习率 0.03 训练 50000 次

隐藏层 : 24

均方误差 :

6.831733978068533E-4

0.08560983689953569

version 10 :

先学习率为 0.05 训练 500000 次, 在学习率 0.03 训练 1000000 次

隐藏层 : 24

均方误差 :

0.0024620292931163902

0.06799154340723068

version 11 :

先学习率为 $0.05/(1+n/25000)$ 训练 100000 次

隐藏层 : 24

均方误差 :

3.834032317765505E-4

0.02993533076738366

2. Sine 调整过程 :

version 1 :

学习率 : 0.08

隐藏层 : 10

次数 : 10000000

均方误差 : 1.7457988739208422E-4

version 2 :

学习率 : 0.05

隐藏层 : 10

次数 : 10000000

均方误差：8.14455874190704E-5
version 3：
学习率：0.03
隐藏层：10
次数：10000000
均方误差：2.7205687386504354E-4
version 4：
学习率：0.05
隐藏层：15
次数：1000000
均方误差：1.4314122929680312E-4
version 5：
学习率：0.05
隐藏层：15
次数：2000000
均方误差：5.355888316620067E-5
version 6：
学习率：0.05
隐藏层：20
次数：2000000
均方误差：1.5507618258257246E-4
version 7：
隐藏层：15
次数：0.05 训练 1000000 次, 0.03 训练 1000000 次, 0.02 训练 1000000 次, 0.01 训练 1000000 次
均方误差：9.09426295163108E-5
version 8：
隐藏层：15
次数：0.05/(1+n/500000)训练 1000000 次
均方误差：7.274930150667549E-7

3. Letter 调整过程：

version 1：
学习率为：0.05
训练次数：20
隐藏层：28
识别率:(上测试集，下验证集)
0.920051413881748
0.880625
version 2：
将训练次数改为 5
性能明显下降,调回 20

识别率:(上测试集, 下验证集)

0.4665809768637532

0.4621875

version 3 :

隐藏层为 56, 训练 20

0.9016709511568124

0.8728125

性能未有提升, 将隐藏层重新调为 28

version 4

隐藏层为 132, 训练 20

0.9249357326478149

0.8940625

version 5

隐藏层为 224, 训练 20 次

0.927892030848329

0.89

version 6 :

隐藏层为 132, 训练 20 次, 学习率 0.03

性能下降, 重新调整学习率为 0.05

0.8538817480719794

0.7978125

version 8:

隐藏层为 132, 训练 20 次, 学习率 0.08

性能下降, 重新调整学习率为 0.05

0.8838817480719794

0.8678125

version 9:

隐藏层为 132, 训练 40 次, 学习率 0.05

0.9232647814910026

0.8915625

version 10 :

隐藏层为 132, 训练 20 次, 学习率随迭代次数下降 $r = 0.05/(1+n/20)$

0.9342748015911038

0.89625