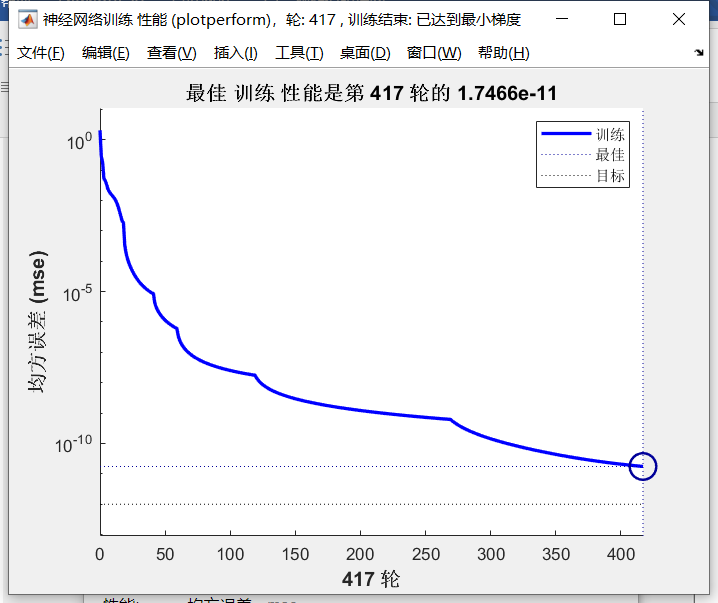
一、当隐含层包含节点数分别为10，12，100，8且激活函数均为‘tansig’时，训练结果如下：



性能如下图所示：

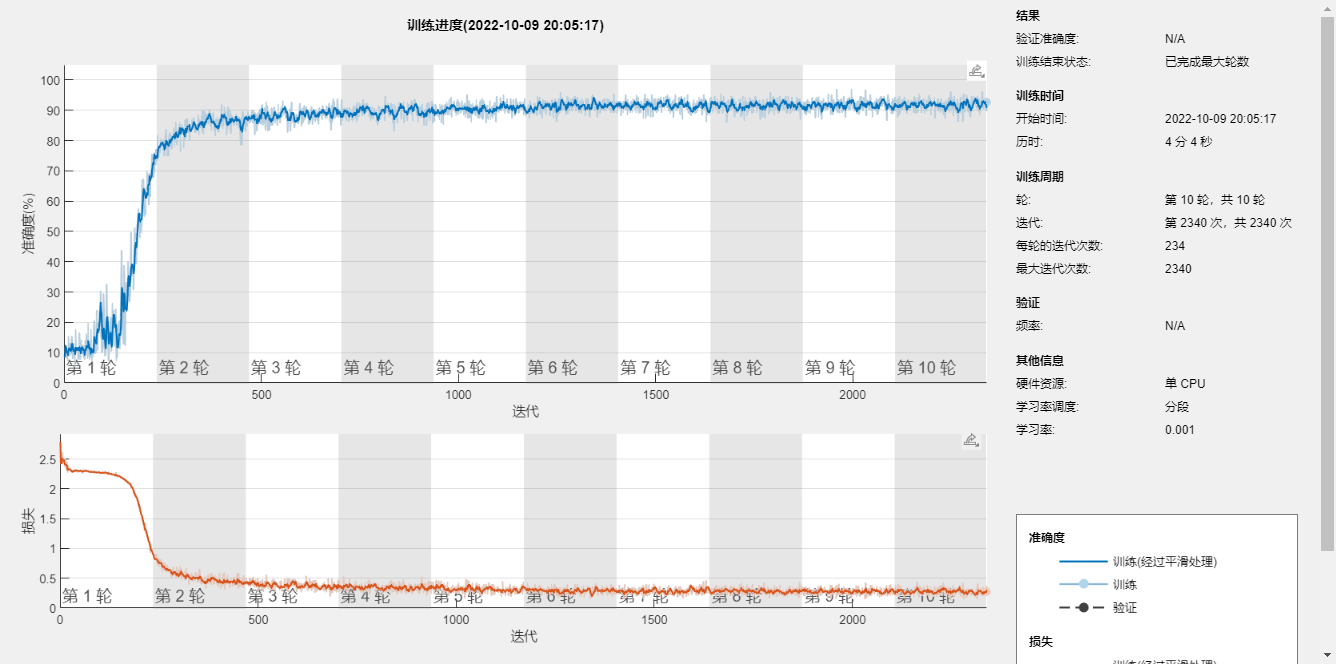


在测试集上准确率为95%。

改变隐含层节点数量和层数，发现当层数一定时（4层）节点数量越大性能越好，但是训练时间也越长；当节点数量一定时，一定范围内层数较大性能较好。

二、在训练中，发现当使用sigmoid函数过多后，会出现异常，训练过程中损失函数与准确率都无较大变化，并维持在效果较差的水平。在反复调试之后发现将全连接层中的sigmoid函数替换成tanh可以避免发生这种异常，原理未知。

训练过程中准确率与损失函数随训练过程变化如下图所示。



最终，在训练集上测试准确率达到92.58%，在测试集上测试准确率达到92.24%。