训练数据集：

数据集为sklearn.datasets 中breast cancer数据集，这是一个二分类任务，数据集的data shape为（569， 30），他、target shape为(569,)

训练任务：

breast cancer数据集的前400个样本用作训练，后169个样本用作测试。

神经网络结构：

所用神经网络为30\*20\*2的多层感知机（MLP），输入层30个节点，隐藏层20个节点，输出为2个类别，激活函数为tanh,。

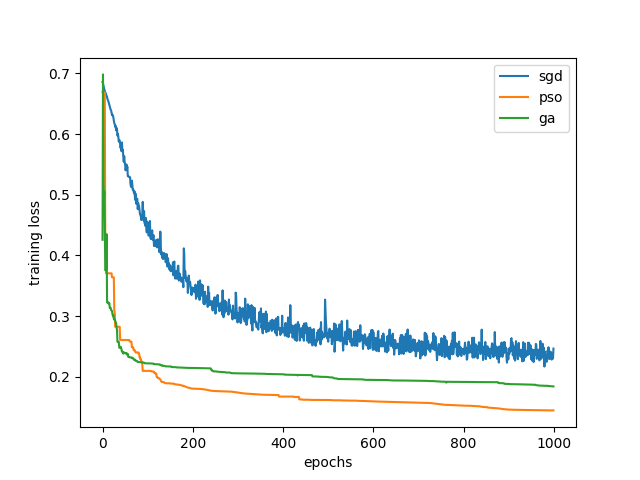
损失函数（优化目标）：

Softmax之后做CrossEntropyLoss（）

训练轮次: 1000

结果：

1. 三种方法的损失函数（优化目标）变化曲线如图所示：



从图中看出，sgd的收敛曲线较为平缓，收敛速度较慢，后期抖动也很明显，并且最终收敛值大于采用pso和ga所得的收敛值。pso和ga的收敛速度都极快，基本上都在100轮以内降到0.2左右，其中pso的收敛速度较ga更快，且收敛的值更小。从损失函数的收敛效果来看，对于breast cancer数据集做二分类，网络的参数量较小，进化算法要优于sgd。

1. 执行时间

三种算法1000轮迭代后的执行时间如表所示，由同一设备的cpu执行运算，可以看出sgd的执行时间明显优于进化算法，两种进化算法的时间差距不大，相较于sgd都比较慢。考虑到训练的数据量很小，而且网络参数量只有662个，sgd和梯度下降的差别还不是很明显，如果参数量很大，sgd可以gpu加速，那么在这三种算法中采用sgd将是唯一的选择。

|  |  |  |  |
| --- | --- | --- | --- |
| algorithms | sgd | pso | ga |
| runtime/s | 14.96 | 44.08 | 55.82 |

1. 测试准确率

三种算法1000轮迭代后的测试准确率如表所示，表中数据显示采用pso时的测试准确率最优，sgd稍次于pso，但差别不大。采用ga时的测试准确率效果要差一些。

|  |  |  |  |
| --- | --- | --- | --- |
| algorithms | sgd | pso | ga |
| test accurracy | 0.927 | 0.941 | 0.875 |

1. 总结

对于数据量小，参数量小的训练任务，进化算法可以快速收敛，可以取得不错的效果，犹以pso效果最好。对于数据量大，参数量大的训练任务，大概只能选择sgd这类的方法。