

计算物理作业 9

王一杰^a

^a 中国科学技术大学

2021 年 12 月 28 日

目录

| | |
|--------------------|----------|
| 1 Problem 1 | 1 |
| 1.1 程序设定 | 1 |
| 1.2 程序实现 | 1 |
| 1.3 结果分析 | 2 |
| 2 Problem 2 | 2 |
| 2.1 程序设定 | 2 |
| 2.2 程序实现 | 3 |
| 2.3 结果分析 | 3 |

1 Problem 1

1.1 程序设定

用二维伊辛模型来描述铁磁相变问题，已知相变发生在温度 $T = 2.26$ 时，温度低于相变点为铁磁性，原子自旋排布相对有序，高于相变点为顺磁性，自旋排布无序。本题考察 40×40 的二维原子点阵，每个原子的自旋向上 ($= 1$) 或向下 ($= 0$)。每个数据点为二维阵列自旋信息，对应标签为铁磁 ($= 1$) 或顺磁 ($= 0$)。利用机器学习随机森林方法来解决这个分类问题。已用蒙特卡罗方法模拟出数据样本，从 $T = 0.25 \rightarrow 4$ 取 16 个温度点，每个温度下有 10000 个数据样本。

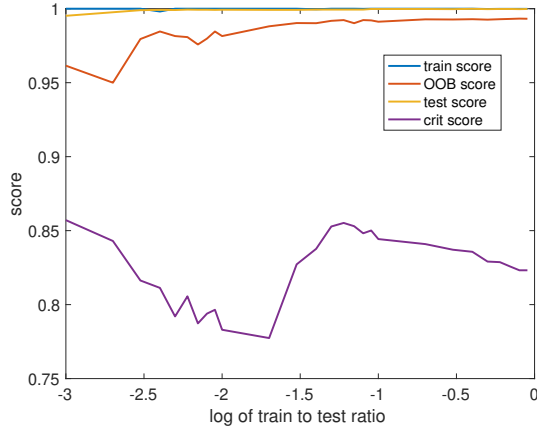
解决如下问题：

- 将测试、训练样本上预言正确率随训练样本占总样本比例的关系画出，训练样本比例区间为 0.1% – 90%；
- 确定训练样本 90%，测试 10%，研究正确率随树木颗数、树木深度，以及叶子上最小样本数的关系，画图。最后给出最优超参选择。

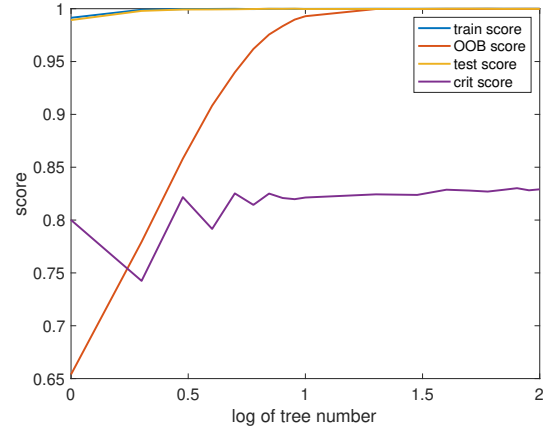
1.2 程序实现

程序实现由参考网站即可。

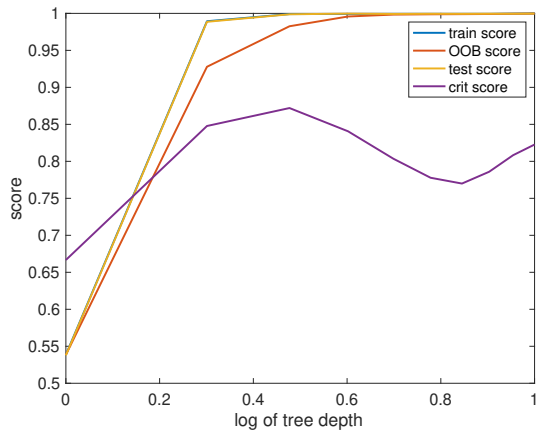
代入参数，得到结果如图 1 所示：



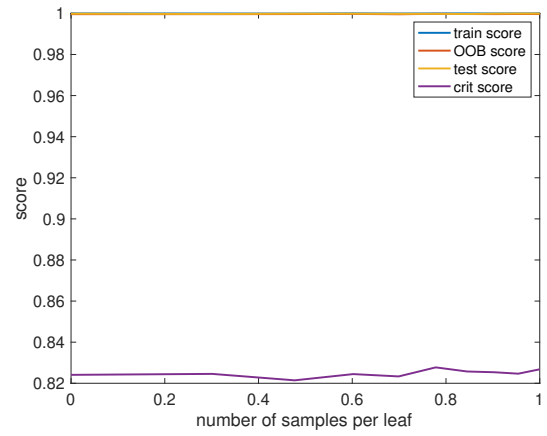
(a) 预言正确率随训练样本占总样本比例的关系



(b) 预言正确率随树木颗数的关系



(c) 预言正确率随树木深度的关系



(d) 预言正确率随叶子上最小样本数的关系

图 1: Problem 1 运行结果

1.3 结果分析

(1) 从图 1(a) 中可以看出, 预言正确率随训练样本占总样本比例的关系在初期的提升比较明显, 后期不明显, 且有一定的震荡。

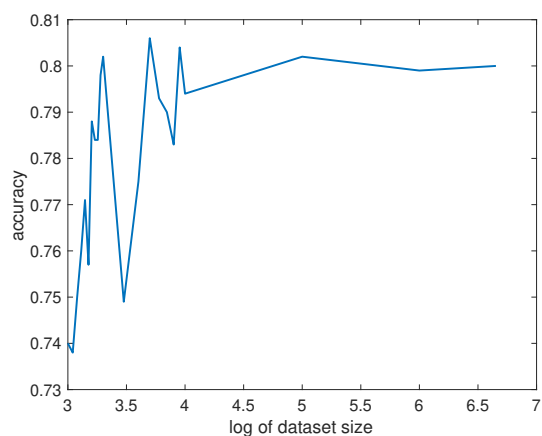
(2) 从图 1(b)(c)(d) 中可以分析一个可能较优的超参为: 树木颗数 $\approx 10^{1.2}$, 树木深度 $\approx 10^{0.5}$, 叶子上最小样本数 $\approx 10^0$

2 Problem 2

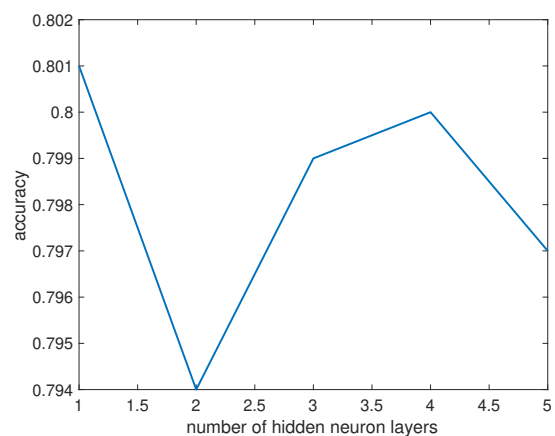
2.1 程序设定

寻找新物理现象是基础物理前沿方向。考察在对撞机实验中寻找超对称现象。利用蒙特卡洛方法产生了共五百万个数据样本, 包含信号和噪声 (即标签)。每个数据点对应实验末态的 18 特征。利用深度神经网络来研究此分类问题。模拟数据样本可从此处下载。建议使用 Pytorch 程序包, 相似例子可见参考书网页。解决如下问题:

- 使用单层隐藏神经元 1000 个, 研究预言正确率与训练样本大小的关系, 训练样本数目范围 $1000 \rightarrow 4500000$, 画出关系图;
- 固定隐藏层神经元每层 100 个, 研究正确率与隐藏层数的关系, 层数范围 1 – 5, 画出关系图。



(a) 预言正确率随训练样本数目的关系



(b) 预言正确率随隐藏神经元层数的关系

图 2: Problem 2 运行结果

2.2 程序实现

程序实现由参考网站即可。

代入参数，得到结果如图 2 所示：

2.3 结果分析

(1) 从图 2(a) 中可以看出随着训练样本数目的增加，预言正确率前期上升很快，后期基本不变，维持在 80% 左右。

(2) 从图 2(b) 中可以看出，隐藏神经元层数在这个问题中和该参数环境下，与正确率基本无关，维持在 80% 左右。