2021 秋 计算物理 B，机器学习作业

练习题 1：

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 x 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→4 取16 个温度点，每个温度下有 10000 个数据样本。模拟数据可由[1]处下载。基于模拟数据开展机器学习（建议使用 sk-

learn 程序包），相似例子可见参考书网页[2]。

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]<https://rec.ustc.edu.cn/share/7ac80610-73a4-11ed-8ae6-b71aa73bf3f7>

[2]<https://physics.bu.edu/~pankajm/ML-Notebooks/HTML/NB9_CVIII-randomforests_ising.html>