



PROJECT SPECIFICATION

Predicting Boston Housing Prices

数据研究

标准	符合规格
统计分析	请求的所有 Boston Housing 数据集统计数据均已得到精确计算。学生可恰当利用 NumPy 功能获得这些结果。
属性观察	学生正确解释各项属性与目标变量增加或减少之间的关联。

开发模型

标准	符合规格
匹配度	学生正确判断假设模型是否能根据其 R^2 分数成功捕捉目标变量的方差。性能指标在代码中正确实施。
训练与测试	学生合理解释为何要为某个模型将数据集分解为训练子集和测试子集。训练和测试分解会在代码中正确实施。

分析模型性能

标准	符合规格
学习数据	随着训练点的不断增加，学生正确判断图表中训练及测试曲线的走向并讨论该模型是否会得益于更多的训练点。
偏差和方差分析	学生提供最大深度为 1 和 10 的分析。如果模型偏差或方差较高，请针对每个图形给出合理的理由。
最佳猜测最佳模型	学生根据合理的理由使用模型复杂度图形选择最佳猜测最佳模型。

评估模型性能

标准	符合规格
网格搜索	学生准确说明网格搜索算法，并简要探讨该算法的用途。GridSearchCV 会在代码中正确实施。

标准	符合规格
交叉验证	学生准确说明如何对模型进行交叉验证，以及它对网格搜索的作用。除非有合理的理由，否则不得对默认的 3 折交叉验证以外的 GridSearchCV 部分进行修改。
拟合模型	学生在代码中正确实施 <code>fit_model</code> 函数。
参数经过调整的最佳模型	学生根据参数调整确定最佳模型，并将此模型与他们选择的模型进行对比。
预测出售价格	学生报告表格所列三位客户的预测出售价格，根据已知数据和先前计算出的描述性统计，讨论这些价格是否合理。
应用	学生深入讨论支持或反对使用他们的模型预测房屋售价的理由。

（可选）预测北京房价

标准	符合规格
代码部分	学生用代码实现了数据分割与重排、训练模型、对测试集进行测试并返回分数。使用交叉验证对参数进行调优并选出最佳参数，比较两者的差别，最终得出最佳模型对测试集的预测分数。
文字答案	学生的回答与其实现的代码相吻合。并表达了自己的观点。

Suggestions to Make Your Project Stand Out!

你可以在[这里查看](#)此评审标准的英文版本。

Student FAQ