

项目

Predicting Boston Housing Prices

此部分属于 Machine Learning Engineer Nanodegree Program

项目审阅

代码审阅

注释

与大家分享你取得的成绩！ 

Meets Specifications

恭喜你通过了这个项目！

同时还给出了一些建议和分享，希望能够对你有所帮助。

加油！继续后面的学习吧～

分析数据



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。



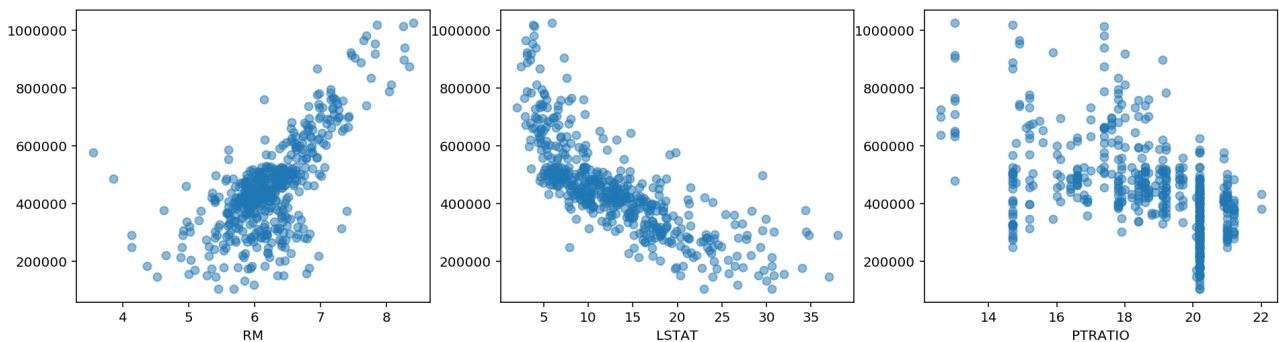
学生正确解释各项属性与目标变量增加或减少之间的关联。

不错的分析，还有一种非常好的思路是通过绘制散点图可视化数据来验证自己的猜想，可以参考下面的代码：

```
# 载入画图所需要的库 matplotlib
import matplotlib.pyplot as plt

# 使输出的图像以更高清的方式显示
%config InlineBackend.figure_format = 'retina'

# 调整图像的宽高
plt.figure(figsize=(16, 4))
for i, key in enumerate(['RM', 'LSTAT', 'PTRATIO']):
    plt.subplot(1, 3, i+1)
    plt.xlabel(key)
    plt.scatter(data[key], data['MEDV'], alpha=0.5)
```



学生合理解释为何要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现。

你可能需要注意这个问题：我们是不能根据测试集来调整模型的，用来调整模型的是验证集。
如果没有测试集我们不能判断模型的拟合程度，无法知道模型泛化能力如何，并不是一定不好，只是无法评价。

模型衡量标准

✓
性能指标在代码中正确实现。（如果做了可选题，但不正确，算过不通过）

非常好，你计算出了正确的 R2 score，它预测得不错。

可汗学院这个[视频](#)很详细的讲解了R2 score的计算过程可以作为参考帮助我们更好的理解R2 score。

✓
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。

分析模型的表现

✓
随着训练点的不断增加，学生正确判断图表中训练集和验证集曲线的走向并讨论该模型是否会得益于更多的训练点。

✓
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。

很好的判断，高偏差通常是由于模型太简单（即模型欠拟合），不能很好的拟合测试集，训练分数、验证分数、测试分数通常都比较低；高方差通常是由于模型过于复杂（即模型过拟合），模型在训练集上表现得很好，在验证集和测试集上得分确比较低，泛化能力差。

✓
学生根据合理的理由使用模型复杂度图形猜测最优模型的参数。

评估模型性能

✓
学生准确说明网格搜索算法，并简要探讨该算法的用途。

非常好！你理解了网格搜索算法。

✓
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。

很好，你的回答中包含了K折交叉验证算法中的关键部分，K折交叉验证避免了因为数据集划分的偶然性造成的评分偏高或偏低的问题。对于每一组参数对应的模型，通过使用不同的训练集和验证集训练然后取K次评分的平均来得到最终成绩来保证评分的客观和准确，从而准确定位到给出参数中的最优参数。

此外，对于引用的部分需要给出引用说明。

✓
学生在代码中正确实现 `fit_model` 函数。（如果做了可选题，但不正确，算过不通过）

很好的实现，你还可以做几点改进：

- 设定KFold的默认参数。比如 `n_splits=10, random_state=1, shuffle=True`。
- 使用KFold和初始化模型的时候设置random state。

✓
学生根据参数调整确定最佳模型，并将此模型的参数与他们猜测的最佳参数进行对比

进行预测

✓
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。

学生计算了最优模型在测试集上的决定系数，并给出了合理的分析。



学生可以合理分析最优模型是否具有健壮性。



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

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。



学生的回答与其实现的代码相吻合。并表达了自己的观点。



 [下载项目](#)

[返回 PATH](#)

给这次审阅打分



[学员 FAQ](#)

[Reviewer Agreement](#)