

项目

Titanic Survival Exploration

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

整体完成的很好，这只是个热身项目，在于熟悉项目提交和jupyter notebook，后面就真正进入机器学习的世界啦～

回答问题

学生运行 `predictions_0` 函数并得到了预测结果。

这部分主要在于帮助你熟悉Python的书写规范和数据导入，初步处理，探索性分型的过程。尽管我们不是一门Python的课程，但是学习代码的编写规范有利于提高代码的可读性和方便调试，在这里我附上Google开源项目风格指南（Python），希望对你有帮助：https://github.com/zh-google-styleguide/zh-google-styleguide/blob/master/google-python-styleguide/python_language_rules.rst

学生正确构建 `predictions_1` 函数。得出了预测准确率。

- 这个部分为分析过程提供了一个范本：通过观察数据特征，依据常识作出合理的推测和假设，利用Python的可视化工具初步证实我们的假设，选择合适的模型，在将我们的想法转换为合适的模型并给予实现。
- 这个分析过程对于建模非常的重要，可以帮助我们减少实际上不相关的特征被错误的加入到模型中，尽管在一些模型里，比如线性回归，在建模后期可以通过一定的方法将这些不相关的特征识别出来，但既然能够通过前期的数据观察排除，何不在一开始就做好呢，有句话在建模领域非常有名：garbage in, garbage out。

学生正确构建 `predictions_2` 函数。得出了预测准确率。

注意我们可以精简我们的代码，请参考下面的代码：

```
if (passenger['Sex'] == 'female' or passenger['Age'] < 10):
    predictions.append(1)
else:
    predictions.append(0)
```

学生正确构建 `predictions_3` 函数，预测结果准确率大于80%。学生记录下了完成函数构建的步骤，包括尝试了那些特征，以及从选取特征到最后完成函数构建的中间步骤。

非常好，你的准确率在我们的学生中是非常高的。其实embarked这个特征也不错，如果和其他特征结合起来，也能得到不错的准确率，你可以自己试试看：

```
if passenger["Sex"] == "female":
    if passenger["Pclass"] == 1 or passenger["Pclass"] == 2:
        predictions.append(1)
else:
    if passenger["Embarked"] == "C" or passenger["Embarked"] == "Q":
        predictions.append(1)
    else:
```

```
        predictions.append(0)

    else:
        if passenger["Age"] < 10.0:
            predictions.append(1)
        else:
            predictions.append(0)
```

希望对你有帮助。

很多特征看似没有包含什么信息，但是实际上可能和一些隐含的特征有强相关关系，所以也能作为纳入模型的变量。再比如“name”这个看似无关的变量，但是其中包含了“Mr”，“Mrs”甚至“Dr”等称呼，这些称呼可能会体现一些信息出来，也能作为有意义的指标。Domain knowledge很重要，但是拿到数据前的EDA（探索式数据分析）也是很重要的。



学生描述了一个监督学习的应用。描述中清晰地描述了需要预测的特征，以及提到了至少两个数据集里的数据特征。

非常好的例子，你举出的特征也非常好，我们在进行监督学习的时候，特征对于我们的模型构建是非常重要的。所以在今后的学习中，你不仅要学会如何实现一个算法，也要关注“特征工程”，这是在现实生活中机器学习工程师经常需要做事情。

 [下载项目](#)

[返回 PATH](#)

给这次审阅打分



[学员 FAQ](#)