

Titanic Survival Exploration

此部分属于 Machine Learning Engineer Nanodegree Program

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。

很好，你实现了你的第一个模型！这里我们做的只是最简单的预测所有乘客不会生还，你可以看到，即使是这样你也获得了61.62%的准确率，这表明我们的数据集是不平衡的，也就是说死亡的乘客比生还的多。我们会在后面的项目中看到，如果数据出现不平衡的现象，是需要用一系列方法去处理的，否则模型的表现会不如预期。



学生正确构建 `predictions_1` 函数。得出了预测准确率。

非常好！你的结果显示了女性比男性有着更高的生存概率，你有没有想过为什么女性比男性更容易存活呢？是因为她们更适应冷水的环境吗，还是因为她们比男性更强壮？这些问题也许看上去无厘头，但是实际上却有实际意义，这在机器学习中称为“domain knowledge”。我们在机器学习的过程中，如果能够有效的运用正确的domain knowledge，可以让我们的机器学习算法更有效。



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。

准确率不错，如果你想把age, sex, pclass等其他特征综合起来提高准确率，你可以参考下下面的代码：

```
if passenger['Sex'] == 'female':
    if passenger['Pclass'] in [1, 2]:
        predictions.append(1)
    else:
        if passenger['Embarked'] in ['C', 'Q']:
            predictions.append(1)
        else:
            predictions.append(0)
elif passenger['Age'] < 10:
    if passenger['SibSp'] in [0, 1]:
        predictions.append(1)
    else:
        predictions.append(0)
else:
    predictions.append(0)
```



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。

很好，这不仅是单纯的机器学习问题，还是自然语言处理的问题

给这次审阅打分

