



PROJECT SPECIFICATION

Finding Donors for CharityML

数据探索

标准	符合规格
数据探索	学生正确地计算了： <ul style="list-style-type: none">记录的数目收入大于50000美金的人数收入小于等于50000美金的人数收入大于50000美金的人数所占百分比

准备数据

标准	符合规格
数据处理	学生正确的实现了对特征和输入数据的 one-hot encoding。

评估模型表现

标准	符合规格
问题一： 简单预测表现	学生正确的计算了简单预测的准确率和F1分数。
问题2： 模型应用	学生解释了选择这几个模型的原因，并说明了每一个模型的优缺点。
创建一个训练和预测的流程	学生成功的实现了一个监督学习算法的流程。
初始模型表现	学生正确的实现了三个监督学习模型，得出了模型表现可视化的图表。

优化结果

标准	符合规格
----	------

问题三： 选择最好的模型	在考虑了计算成本、模型表现和数据特点之后，学生选出了最好的模型并给出了充足的理由。
问题四： 用普通人能听懂的话来描述模型	学生能够用清晰简洁的话来向一个没有机器学习或任何其他技术背景的人来解释最优模型的工作原理。
模型调优	最终模型利用了网格搜索进行参数调优，至少挑战了一个参数，并且至少有三个可选值。如果模型参数不需要任何调整，学生需要给出明确的理由。
问题五： 最终模型评估	学生在表格中正确汇报了调优过后、调优之前以及基准模型的准确率和 F1 分数。学生把最终模型的结果与之前得到的结果进行了对比。

特征的重要性

标准	符合规格
问题六： 观察特征相关性	学生列出了他们认为对预测个人收入最重要的5个特征，同时给出了选择这些特征的理由。
问题七： 提取特征重要性	学生调用了一个监督学习模型的 <code>feature_importances_</code> 属性。此外，学生列出了这些重要的特征并讨论了这些特征的相同点和不同点。
问题八： 特征选择效果	学生用最重要的5个特征建模并分析了和对比了改模型与问题五中的最优模型的表现。

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](#)