

课程设计

利用kaggle中titanic的train数据构建对survival情况的预测模型，并对test数据集进行预测。

要求：

1. 分组完成，每组2-3人；
2. 使用任意一种分类模型；
3. 需要包含数据预处理过程；
4. 给出对应的可视化结果；
5. 训练过程中要包含cross validation的步骤；
6. 给出训练和预测过程的分类指标相应结果；
7. 程序可以使用任意语言；
8. 最终需要提交程序和相应的报告；
9. 报告中需列出每名成员具体的工作内容及工作所占整个工程的百分比。

数据网址：<https://www.kaggle.com/c/titanic/data>