

任务

为了便于考察目的，我们为您准备好了数据集(demo_data.tar.gz)，这是一个结构化数据集，具体包括了10000个数据点，每个数据点包括300多个特征和一个真实的y值。

请在给定的时间内，完成下述任务：

1. 请对特征数据进行探索性数据分析(EDA), 并给出您对这些特征数据定量/定性的结论
2. 将上述数据作为监督学习的样本数据, 请建立回归/分类模型(自选一个即可), 并对模型性能进行定量评估.

最后请将您的代码和结果以notebook (ipynb)的形式提交, 如有使用到了外部文件, 请一并附上.

由于存在时间和空间的约束, 我们建议您重点关注下述几个角度, 这样可以更为高效高质的完成这个任务:

- 好的代码风格, 和优美的字体一样让人赏心悦目
- 一图胜千言
- 整体分析框架的完备性
- 确保关键环节没有遗漏
- 定性分析也很重要, 但这应该在定量分析的基础上完成

请务必理解: 我们所考察的并不是某个具体指标的高或低, 而是在这个过程中体现出的能力和潜力.