

数据集介绍

本次作业的两道编程题都使用同样一份数据集，保存在 `data.csv` 中。

此数据为某公司用户在第三方交易平台的用户的日常刷卡行为数据。数据提供了一批用户行为记录，包括：年龄、银行卡数、借贷比率、交易笔数、所有行为均值、所有行为最大值等。

数据集共有 27 列，其中前 26 列为收集到的特征，最后一列(black)为是否违约；共含有 7800 条数据，我们随机划分了 80% 作为训练集，其余为测试集。

任务说明

考虑到同学们的知识背景与编程能力存在一定差异，本次作业我们提供了较为详细的代码框架，只需要在算法的主体部分补全代码块即可实现需要的功能。

其中，每一处必须补全的代码空缺都通过 `TODO` 标注（第三题有 3 处、第五题有 4 处）。你应当删除代码块中占位使用的 `assert 0` 语句，并补全必要的代码。

在作业中，你可以使用 `numpy`, `pandas` 等矩阵运算包，但不可以使用 `autograd`, `cvxopt` 等自动求导或凸优化求解包。此外，除非题目中特殊标注，否则不要使用 `sklearn` 中的方法。