

机器学习实验 - 2023 春

实验一：线性模型

实验内容：线性回归模型，逻辑回归模型

作业提交截止时间：2023 /03/17 23:59:59

环境要求：

Python, numpy 支持多维度的数组和矩阵运算, pandas 数据处理和分析工具, Matplotlib 图形化工具, sklearn 机器学习库。

任务一：线性回归模型

本任务中你将使用一元线性回归来预测食厅的利润。假设你是一家特许餐厅的首席执行官，正在考虑在不同的城市开设一家新的分店。该连锁店已经在不同的城市有分店，你有这些城市的**利润**和**人口**数据。你希望使用这些数据来帮助您选择下一个要扩展到的城市。

文件 `ex1data.csv` 包含我们的线性回归问题的数据集。`Population` 代表**一个城市的人口**，`profit` 代表**这个城市的餐厅利润**。利润的**负值表示亏损**。

请将 70%的数据用作训练集，30%的数据用作测试集，使用留出法对以上模型进行验证。

任务二：逻辑回归模型

本任务中你将建立一个**逻辑回归模型**来预测一个学生是否被大学录取。假设你是一所大学系的管理员，你想**根据两次考试的成绩**来决定每个申请人的录取机会。你有以前申请者的历史数据，可以用作逻辑回归的训练集。对于每个培训示例，你都有申请人在两次考试中的分数和录取决定。你的任务是建立一个分类模型，根据这两次考试的分数来**估计申请人的录取概率**。

文件 `ex1data2.csv` 包含我们的逻辑回归问题的数据集，学生的两门成绩 `Exam1`，`Exam2` 和是否被录取 `Accepted` (1 为录取，0 为未录取)，请利用逻辑回归知识根据学生成绩判断学生是否会被录取，并将结果可视化。

请使用 **5 折交叉验证法**对以上模型进行验证。

作业提交内容：

1. 作业代码
2. 实验结果及分析（PDF 格式）
3. 请将以上文件打包并以 zip 格式压缩，命名为 学号+姓名.zip，邮件发送到 facanhe@163.com