

机器学习任务 1

(1) 熟悉 PCA 和线性回归算法。

(2) 完成教程 1-5 中介绍的编程任务。

(3) 自己编写一个线性回归算法。

- 使用波士顿房屋数据集（506 个样本和 13 个特征变量），利用给定的特征预测房屋价格的价值。

- 将输入数据分成训练集和测试集（测试集包括 10% 的样本）。

- 算法的输出应包括从整个训练集学习到的模型、通过 10 倍交叉验证得出的训练集平均均方误差（MSE）。对于脊回归模型，通过 10 倍交叉验证确定超参数值。

(4) 将实验过程和结果写成实验报告。(内容严谨，图文并茂)

。

- 查看数据中是否有缺失值。

- 探索性数据分析（例如，绘制目标变量的分布图、创建衡量变量间线性关系的相关矩阵并绘制相关矩阵图、显示测试集的均方误差 (MSE)、绘制原始房价与预测房价之间的散点图）。

(5) 源代码应作为附录提交（只需提交回归算法的源代码）。