2020秋季学期 机器学习-第一次作业

**线性回归 (linear regression)：**为销售某商品，分别向电视(TV)，广播(radio)和报纸(newspaper)三种媒体投放广告。数据集描述了给定三种媒体广告投放预算下的商品销售量(sales)。请对数据集使用10折交叉验证，训练一个线性回归模型，预测商品销售量与三种媒体广告预算之间的关系。作业提交模型训练代码和报告，报告回答以下问题：

（1）描述10折交叉验证对数据集的处理；（5分）

（2）描述所使用的线性模型；（10分）

（3）描述训练模型所使用的算法；（15分）

（4）分析模型训练结果，包括训练误差和测试误差；（10分）

（5）总结模型训练过程中的收获。（10分）

报告得分规则：写出具体的数学公式及推导过程。

代码得分规则（共50分）：不限定代码语言（若用python，请提交Jupyter notebook文件）。用测试数据集测试每个学习小组提交的模型的测试误差并排序，测试误差最小的学习小组代码得分50分，其余小组按排序依次减1分。

注：代码或报告雷同将严重影响最后得分。

提交时间：**2020年11月8日22:00之前**

作业格式：代码+报告（报告为pdf文件，报告首页标明小组成员姓名+学号）

作业命名（邮件名称和提交作业文件名称）：学习小组xx-机器学习-第一次作业，例如：学习小组12-机器学习-第一次作业

邮件形式发送至邮箱：zhourm3[at]mail2.sysu.edu.cn