

深度学习导论作业 1-1

数据集说明

使用波士顿房价数据集，包含 506 个样本，13 个特征与 1 个连续标签（房价中位数 MEDV）

关键特征示例：

RM（房间数）、CRIM（犯罪率）、LSTAT（低收入人群比例）等

实验步骤

加载数据集(加载方法参考 load_dataset.py)

将数据集随机划分为训练集和测试集(可以按照 5:1 或 4:1 的比例划分)

将训练集进一步划分为训练/验证

构建前馈神经网络，输入层维度需与特征数（13）匹配，输出层为 1 个神经元（回归任务）。

使用均方误差（MSE）作为损失函数，优化器可选择 SGD 或 Adam。

每个 epoch 输出训练集和验证集的损失，批量大小（batch size）建议设为 16 或 32。

保存最佳模型参数（基于验证集损失）。

在测试集上评估模型性能。

实验分析

网络深度的影响

实验设置：

固定学习率（如 0.001）、激活函数（如 ReLU），设计 3 种结构(不强制是这三种，可以自己设计)：

浅层网络：1 个隐藏层

中等网络：2 个隐藏层

深层网络：3 个隐藏层

分析方向：

训练/验证/测试损失随深度增加的变化趋势。

深层网络是否出现过拟合（训练损失低但验证/测试损失高）。

学习率的影响

实验设置：

固定网络结构（如 2 层隐藏层）、激活函数（如 ReLU），测试学习率：0.1、0.01、0.001、0.0001。（不强制，可以自己设置）

分析方向：

不同学习率下训练损失的收敛速度与稳定性。（可以通过画出训练/验证/测试 loss 曲线分析）

学习率过大（损失震荡/不收敛）或过小（收敛缓慢）的表现。

激活函数的影响

实验设置：

固定网络结构（2 层隐藏层）、学习率（0.001），测试激活函数：Sigmoid、Tanh、ReLU。

分析方向：

不同激活函数对训练效率的影响（如 ReLU 是否收敛更快）。

作业提交方式

将代码和实验报告打包成压缩包，命名为'学号-姓名-hw1'的形式在 bb 系统上提交，不需要提交模型权重