

# “深度学习和计算机视觉（实践）”课程实验一： 使用梯度下降训练线性模型

廖振宇

日期：2021 年 9 月 22 日

## 1 实验背景

**课程内容回顾** 在理论课程中，我们回顾了机器学习的基本概念，模型的评估和选择，线性模型和广义线性模型的概念和相关的梯度下降的优化方法。

**要求** 不使用 Tensorflow、Pytorch 或者 scikit-learn 等机器学习框架，仅使用 Numpy，Scipy 和 Matplotlib 等 Python 常用科学计算和可视化库，使用梯度下降训练线性模型，实现（两类）正态分布数据的分类。

## 2 实验要求

- 人工生成数据，分类问题：两类正态分布  $\mathcal{N}(\pm\mu, \mathbf{I}_d)$ ， $d = 2$ ， $\mu = [1; -1]$ ，数据总数目  $m = 500$
- 选取训练和测试集
- 构建线性分类模型，推导梯度公式，手写代码更新梯度
- 使用梯度下降的方法训练模型，针对得到的不同模型（不同初始化方法、不同梯度下降步长、不同训练步数）进行模型评估和选择；尤其的，将线性回归的解析解同梯度下降得到的解比较，有什么结论？
- （选做）保持其余条件不变，更改  $d = 500$ ，选取  $\mu = [1; \mathbf{0}_{d-1}]$ ，完成同样的实验，画出测试误差随训练步数的变化情况，和之前  $d = 2$  的情形进行比较，是否观察到了更加严重的过拟合现象？类似的，选取  $d > 500$ ，比如  $d = 1000$ ，可以观察到什么现象？
- PPT 汇报（每组 3min），提交 2-4 页实验报告，需简要叙述方法原理、实验步骤、方法参数讨论、实验结果；需明确说明组员分工、给出组内排名（可标注同等贡献 #）。