

# 机器学习课程设计安排

2019.12.19-

# 课程设计目标及要求

- 注重数学建模基础，通过课程设计深入掌握机器学习中基本原理和基本算法。希望从多个维度考核驱动，锻炼同学们的动手编码能力、文档撰写能力、表达能力、组织能力、合作与自学能力等。
- 课程设计建议以团队讨论、分工合作的形式完成，并进行汇报和提交。
- 课程设计报告的内容组织要求按照软件工程文档规范，覆盖从问题需求分析，机器模型概要设计，算法详细设计，算法实现，实验分析和优化等各个方面内容。
- 考核评价依据：实验成果（完成度、完成质量），汇报答辩情况（能否在规定时间内以清晰逻辑讲清楚小组的工作、收获和不足，代码演示录屏效果等），实验报告（规范性、完整性、原创性等）

# 时间安排

- 上午：8:30-11:30 下午：14:00-17:00
- 12.19周六上午/下午
  - 实验一：编程实现AdaBoost模型，不剪枝决策树为基学习器，在西瓜数据集3.0a上训练一个AdaBoost集成，并与教材图8.4进行比较。
  - 实验二：编程实现Bagging模型，以决策树桩为基学习器，在西瓜数据集3.0a上训练一个Bagging集成，并与教材图8.6进行比较。
- 12.20周日上午/下午
  - 实验三：设计一个BP改进算法，能够动态调整学习率显著提升收敛速度，选择两个UCI数据集与累积BP算法进行实验比较。
  - 实验四：编程实现基于SMO算法的SVM模型，选择两个UCI数据集进行实验及测试。
- 12.21周一上午/下午
  - 实验五：设计并实现一种改进的朴素贝叶斯分类器，选择两个UCI数据集进行实验及测试。  
(改进方案任选其一：结构扩展，属性加权，实例加权，属性选择等)

# 时间安排

- 12.22周二上午/下午
  - 实验六：编程实现KNN分类器，选择两个UCI数据集进行实验及测试。
  - 实验七：编程实现C4.5决策树算法，选择两个UCI数据集进行实验及测试。
- 12.23周三上午/下午
  - 实验八：自学XGBoost的设计原理，并从网上下载开源代码，选择两个UCI数据集进行实验及测试。
  - 实验九：自学Sklearn的PCA函数使用方法，对Yale人脸数据集进行降维，观察并分析前20个特征向量对应的图像。
- 12.24周四上午/下午 分组答辩汇报
- 12.25周五上午/下午 分组答辩汇报