



中山大學  
SUN YAT-SEN UNIVERSITY

# 机器学习与数据挖掘课堂作业

## 集成学习

数据科学与计算机学院 17大数据与人工智能

17341015 陈鸿峥

**问题 1.** By combining bagging and boosting strategies, design a more effective ensemble framework.

**解答.** 由于bagging过程中会对训练集进行划分，然后训练每一基学习器，而每一基学习器都是相互独立的，因此可以用boosting的方法训练bagging的基学习器。当boosting训练出的基学习器足够强时（比如比单一的决策树要好），那利用bagging将这些基学习器整合在一起，性能也会更优。又由于bagging中每个基学习器的训练都可以并行完成，因此即使在基学习器中使用了boosting，依然可以保持较快的速度。