

集成学习(Ensembles Learning)思考题

- 一. 什么是集成学习?集成学习的原理是什么?
- 二. 集成学习对数据 (特征) 或模型的要求是什么?
- 三. 下面集成方法的原理是什么?如何对数据 (特征) 进行抽样/选择? 如何对模型的结果进行组合?
 1. 基本集成方法
 - 1.1 最大投票 (Max Voting)
 - 1.2 平均 (Averaging)
 - 1.3 加权平均 (Weighted averaging)
 - 1.4 包裹法 (Wrapping up)
 2. 高级集成技术
 - 2.1 装袋(Bagging)
 - 2.2 提高(Boosting),
 - 2.3 堆垛(Stacking)
 - 2.4 混合(Blending)
 3. 基于装袋和提高的算法
 - 3.1 装袋元估计(Bagging meta-estimator)
 - 3.2 随机森林(Random Forest)
 - 3.3 AdaBoost
 - 3.4 GBM
 - 3.5 XGB
 - 3.6 Light GBM

3.7 CatBoost

四. 集成学习的优势和缺陷。