

Домашнее задание №5 по курсу «Машинное обучение»: Неравномерная изучаемость. Выбор модели

Колесов Алексей

3 октября 2018 г.

1 Задания

1. В алгоритм k -fold на что влияет параметр k ? В каких случаях стоит выбирать его большим, в каких маленьким?
2. Почему в практических задачах делят выборку на три части: тренировочную, валидационную и тестовую? Почему не хватает первых двух?
3. Рассмотрим альтернативу алгоритму k -fold — алгоритм «leave-all-out». В этом алгоритме мы разбиваем всеми возможными способами выборку на две части — на одной части тренируем алгоритм, а на другой (отложенной) измеряем качество. Затем все полученные измерения усредняем и используем в качестве оценки L_D . В чём преимущества и недостатки данного подхода по сравнению с k -fold?