



ТЕСТ МОЖНО СДАТЬ ТОЛЬКО 1 РАЗ, НАЖАВ НА КНОПКУ "Сохранить решение"

В вопросе может быть **несколько** вариантов правильного ответа (то есть от 1 до кол-ва ответов в вопросе). Вопрос засчитывается, если выбраны **ВСЕ** правильные варианты и **НЕ ВЫБРАНЫ** **ВСЕ** неправильные варианты.

В тестовых заданиях первая галочка — правильный ответ, вторая галочка — выбранный ответ. Цвет обозначает, правильно ли в данном пункте поставлена галочка. Если все пункты верные (галочки совпадают / все пункты зеленые), то за задание ставится полный балл, в противном случае ставится 0 баллов.

1. Выберите верное утверждение относительно алгоритма настройки (обучения) дерева CART:

- ☐ ☐ настройка выбирает максимально экономичное описание данных бинарным деревом из всех возможных в терминах общего числа вершин
- ☐ ☐ настройка выбирает максимально экономичное описание данных бинарным деревом из всех возможных в терминах общего числа листьев (терминальных вершин)
- ☒ ☐ алгоритм настраивает условие для каждой внутренней вершины локально оптимально с точки зрения изменения критерия информативности лишь на один шаг вперед, без глобальных гарантий оптимальности
- ☐ ☒ алгоритм настраивает условие для каждой внутренней вершины локально оптимально с точки зрения изменения критерия информативности на K шагов вперед, без глобальных гарантий оптимальности (K может быть равным 1,2,3,... - параметр метода)
- ☐ ☐ настройка выбирает максимально экономичное описание данных бинарным деревом из всех возможных в терминах глубины дерева

Балл: 0

Комментарий к правильному ответу:

2. В алгоритме бустинга над решающими деревьями деревья нужно брать

- ☒ ☒ малой глубины
- ☐ ☐ большой глубины

Балл: 2

Комментарий к правильному ответу:

3. Высокое значение разброса при малом значении смещения в разложении на смещение и разброс (bias-variance decomposition) свидетельствует, что модель в общем случае будет показывать средние потери



- ☐ ☐ высокие на обучающей и высокие на тестовой выборке
- ☐ ☐ высокие на обучающей, но низкие на тестовой выборке
- ☒ ☒ низкие на обучающей и высокие на тестовой выборке
- ☐ ☐ низкие на обучающей и низкие на тестовой выборке

Балл: 2.0

Комментарий к правильному ответу:

4. Высокое значение смещения при малом значении разброса в разложении на смещение и разброс (bias-variance decomposition) свидетельствует, что модель в общем случае будет показывать средние потери

- ☒ ☒ высокие на обучающей и высокие на тестовой выборке
- ☐ ☐ низкие на обучающей и высокие на тестовой выборке
- ☐ ☐ высокие на обучающей, но низкие на тестовой выборке
- ☐ ☐ низкие на обучающей и низкие на тестовой выборке

Балл: 2.0

Комментарий к правильному ответу:

5. В алгоритме AdaBoost веса объектов, неправильно классифицированные текущим базовым алгоритмом,

- ☐ ☐ уменьшаются
- ☐ ☐ не изменяются
- ☒ ☒ увеличиваются

Балл: 2.0

Комментарий к правильному ответу:

6. Рассмотрим задачу регрессии. Верно ли утверждение, что существуют сколь угодно не точные модели, при усреднении которых можно получить сколь угодно точную?

- ☐ ☐ нет



☒ ☒ да

Балл: 2.0

Комментарий к правильному ответу:

$f1(x)=f(x)+\epsilon$, $f2(x)=f(x)-\epsilon$

7. В алгоритме бустинга базовые алгоритмы строятся

☐ ☐ одновременно, каждый алгоритм зависит от прогнозов всех остальных

☒ ☒ последовательно, следующий алгоритм зависит от прогнозов предыдущих

☐ ☐ одновременно, независимо друг от друга

Балл: 2.0

Комментарий к правильному ответу:

8. Пусть все признаки значимо влияют на отклик. Сравним метод случайных подпространств над решающими деревьями с алгоритмом случайного леса. Базовые алгоритмы какого метода будут зависеть в общем случае (при достаточно глубоких деревьях) от всех признаков (если все признаки значимы)

☒ ☐ у метода случайного леса

☐ ☒ у метода случайных подпространств

Балл: 0

Комментарий к правильному ответу: