



## 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. В алгоритме бустинга над решающими деревьями деревья нужно брать

☒ ☒ малой глубины

☐ ☐ большой глубины

**Балл: 2**

**Комментарий к правильному ответу:**

2. Рассмотрим задачу бинарной классификации с помощью  $M$  произвольных базовых классификаторов, прогнозы которых агрегируются композицией, реализуемой через назначение самого распространенного класса среди прогнозов базовых моделей (majority voting). Верно ли утверждение, что если вероятность ошибки каждого классификатора всегда равна одной и той же константе, строго меньшей 0.5 и ошибки классификаторов происходят независимо друг от друга (независимые случайные величины), то с ростом  $M$  можно сделать вероятность ошибки композицией сколь угодно близкой к нулю?

☒ ☐ да

☐ ☒ нет

**Балл: 0**

**Комментарий к правильному ответу:**

3. 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, но низкие на тестовой выборке



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

**Балл: 2.0**

**Комментарий к правильному ответу:**

4. Выберите функцию потерь от отступа  $M$ , на которую настраивается алгоритм AdaBoost:

- ☐ ☐  $\max\{0, -M\}$
- ☐ ☐  $\max\{0, 1 - M\}$
- ☒ ☒  $e^{-M}$
- ☐ ☐  $\log_2(1 + e^{-M})$

**Балл: 2.0**

**Комментарий к правильному ответу:**

5. Метод случайных подпространств основан на усреднении различных базовых моделях за счет

- ☐ ☐ обучения на разных псевдовыборках, полученных из исходной путем сэмплирования объектов
- ☐ ☒ сэмплирования начальных условий при настройке базовых моделей
- ☐ ☐ сэмплирования гиперпараметров базовых моделей
- ☒ ☐ обучения на разных псевдовыборках, полученных из исходной путем сэмплирования признаков

**Балл: 0**

**Комментарий к правильному ответу:**

6. Можно достичь высокой точности неглубоким деревом CART в случае, если реальная граница между классами:

- ☐ ☐ описывается сферой вокруг одного из объектов
- ☒ ☒ параллельна одной из осей координат
- ☐ ☐ линейна, даже если не параллельна ни одной из осей координат



**Балл: 2.0**

**Комментарий к правильному ответу:**

7. Оценка качества моделей out-of-bag доступна для

- ☒ ☐ комбинации бэггинга и метода случайных подпространств
- ☐ ☐ метода случайных подпространств
- ☐ ☐ усреднения любых базовых алгоритмов
- ☒ ☒ метода бэггинга

**Балл: 0**

**Комментарий к правильному ответу:**

8. Преимущество усреднения прогнозов с неравномерными подбираемыми весами по сравнению с усреднением прогнозов с фиксированными равномерными весами будет в общем случае больше, если

- ☐ ☐ базовые модели примерно похожи по точности
- ☒ ☒ точность базовых моделей сильно различается

**Балл: 2.0**

**Комментарий к правильному ответу:**