

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

- Решайте тест того потока, лектора которого вы слушаете. Скорее всего, ваш выбор будет совпадать с вашим текущим потоком обучения, однако если вы преимущественно ходите/слушаете в записи другого лектора, то выбирайте соответствующий ему поток
- В дальнейшем вам автоматически будет засчитываться только тот поток, который вы выбрали при решении Теста №1 (или первого из решенных вами тестов по времени). Решать тесты других потоков не будет иметь смысла
- Поменять поток можно будет только при индивидуальном обращении к Ане @liebeapp с причиной, почему хотите поменять поток

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. Суть метода ранней остановки обучения для градиентного бустинга заключается в том, чтобы:

- ☐ ☐ Прекращать обучение если качество на проверочной выборке перестало падать
- ☐ ☐ Прекращать обучение если качество на тренировочной выборке перестало падать
- ☐ ☐ Не переобучиться, варьируя сложность ансамбля случайным образом
- ☒ ☐ Не переобучиться, варьируя сложность ансамбля от простого к сложному
- ☐ ☐ Динамически определять размер базовых моделей ансамбля
- ☒ ☐ Динамически определять размер ансамбля
- ☒ ☒ Прекращать обучение если качество на проверочной выборке перестало расти
- ☐ ☐ Не переобучиться, варьируя сложность ансамбля от сложного к простому
- ☒ ☐ Прекращать обучение если качество на тренировочной выборке перестало расти

Балл: 0

Комментарий к правильному ответу:

2. В алгоритме AdaBoost для бинарной классификации каждый базовый алгоритм обучается так, чтобы на тренировочной выборке:

- ☒ ☐ максимизировать взвешенную точность



- ☐ ☐ минимизировать взвешенную квадратичную функцию потерь
- ☐ ☒ минимизировать взвешенную экспоненциальную функцию потерь

Балл: 0

Комментарий к правильному ответу:

3. Для каких из перечисленных ниже типов базовых моделей имеет смысл применять бэггинг:

- ☒ ☒ с маленьким смещением и большой дисперсией
- ☐ ☐ с маленьким смещением и маленькой дисперсией
- ☐ ☐ с большим смещением и большой дисперсией
- ☐ ☐ с большим смещением и маленькой дисперсией

Балл: 1

Комментарий к правильному ответу:

4. В алгоритме AdaBoost для бинарной классификации при выводе формулы пересчета весов базовых классификаторов существенным требованием к базовому классификатору является, чтобы множество значений прогноза отклика принадлежало множеству:

- ☐ ☐ отрезок $[0, 1]$
- ☒ ☒ два значения $\{-1, 1\}$
- ☐ ☐ отрезок $[-1, 1]$
- ☐ ☐ два значения $\{0, 1\}$

Балл: 1

Комментарий к правильному ответу:

5. Основные потенциальные недостатки блендинга (простого стекинга) с отложенной выборкой для обучения мета-алгоритма являются:

- ☒ ☒ исключение части выборки из обучения базовых алгоритмов
- ☐ ☐ недообученность за счет того, что мета-алгоритм учится на тех же примерах, что и базовые алгоритмы



- ☒ ☒ исключение части выборки из обучения мета-алгоритма
- ☐ ☐ переобученность за счет того, что мета-алгоритм учится на тех же примерах, что и базовые алгоритмы

Балл: 1

Комментарий к правильному ответу:

6. Если базовые алгоритмы - линейные бинарные классификаторы (строят разделяющие гиперплоскости в пространстве признаков), то с помощью каких типов ансамблей можно построить много-классовую модель классификации - ансамбль:

- ☒ ☒ ECOC
- ☐ ☒ Бустинг
- ☐ ☐ Бэггинг
- ☐ ☐ усреднение по подвыборкам из кросс-валидации
- ☐ ☒ Смесь экспертов
- ☒ ☒ Стекинг

Балл: 0

Комментарий к правильному ответу:

7. Ньютоновский бустинг, предложенный в XGBoost, позволяет использовать:

- ☐ ☒ Новый алгоритм формирования гипотез (допустимых точек) для поиска разбиения в базовых деревьях на основе первых и вторых производных функции потерь и параметров регуляризации всего ансамбля
- ☐ ☐ L0 - регуляризацию прогнозов базовых деревьев для борьбы с переобучением
- ☐ ☐ Произвольные базовые алгоритмы, не обязательно деревья решений
- ☒ ☒ L2 - регуляризацию прогнозов базовых деревьев для борьбы с переобучением
- ☒ ☒ Новый критерий выбора оптимального разбиения в базовых деревьях на основе первых и вторых производных функции потерь и параметров регуляризации всего ансамбля
- ☐ ☐ Произвольные функции потерь, не обязательно гладкие и дифференцируемые



☒ ☒ L0 - штраф по числу листьев базовых деревьев для борьбы с переобучением

Балл: 0

Комментарий к правильному ответу:

8. Почему градиентный бустинг называется градиентным?

- ☒ ☐ Потому что рассчитывает вес (степень учета) базового классификатора как оценку шага градиентного спуска
- ☐ ☐ Потому что использует антиградиент функции потерь базовых моделей для расчета весов примеров в выборке
- ☒ ☒ Потому что использует антиградиент функции потерь ансамбля в качестве откликов для обучения базовых моделей
- ☐ ☐ Потому что использует градиентный спуск как алгоритм обучения базовых моделей в ансамбле

Балл: 0

Комментарий к правильному ответу:

9. Основные потенциальные недостатки блендинга (простого стекинга) без отложенной выборки для обучения мета-алгоритма являются:

- ☒ ☒ переобученность за счет того, что мета-алгоритм учится на тех же примерах, что и базовые алгоритмы
- ☐ ☐ исключение части выборки из обучения базовых алгоритмов
- ☐ ☐ недообученность за счет того, что мета-алгоритм учится на тех же примерах, что и базовые алгоритмы
- ☐ ☐ исключение части выборки из обучения мета-алгоритма

Балл: 1

Комментарий к правильному ответу:

10. Какие из перечисленных ниже свойств деревьев решений являются наиболее полезными с точки зрения построения бэггинг ансамбля (выберете два наиболее важных):

- ☒ ☒ нелинейность моделируемой зависимости отклика от предикторов
- ☐ ☐ возможность работы с категориальными признаками без дополнительной предобработки данных



- ☐ ☐ возможность работы с пропусками без дополнительной предобработки данных
- ☒ ☒ нестабильность архитектуры от выборки
- ☐ ☐ интерпретируемость

Балл: 1

Комментарий к правильному ответу: