

MADE: Machine Learning. HA N°1.

Deadline: 19/10/2020 23:59 Moscow Time.

Базовые концепты машинного обучения

Выберете из списка задачи, относящиеся к классификации

1 балл

- ☐ Прогнозирование оттока: для каждого клиента требуется определить, с какой вероятностью он покинет компанию через выбранный промежуток времени в будущем
- ☐ Прогноз спроса на товары и услуги: для каждого товара/услуги требуется спрогнозировать спрос с фиксированным горизонтом в будущем
- ☐ Сегментация пользователей: требуется проанализировать аудиторию и разделить её на сегменты (описание сегментов и их количество заранее неизвсетны)
- ☐ Товарные рекомендации: для каждого пользователя сайта магазина техники требуется построить персональные рекомендации товаров магазина
- ☐ Прогноз спроса на наличные в ATM банка: для каждого ATM требуется спрогнозировать сколько наличных будет запрошено пользователями на фиксированном горизонте в будущем
- ☐ Автоматическое определение жанра фильма по текстовой аннотации
- ☐ Анализ эмоциональной окраски сообщений в службе поддержки: определить носит ли сообщение позитивный, негативный или нейтральный характер



Выберете задачи из списка, относящиеся к регрессии

1 балл

- ☐ Прогнозирование оттока: для каждого клиента требуется определить, с какой вероятностью он покинет компанию через выбранный промежуток времени в будущем
- ☐ Прогноз спроса на товары и услуги: для каждого товара/услуги требуется спрогнозировать спрос с фиксированным горизонтом в будущем
- ☐ Сегментация пользователей: требуется проанализировать аудиторию и разделить её на сегменты (описание сегментов и их количество заранее неизвестны)
- ☐ Детектирование производственного брака: требуется на ранней стадии производства выявить продукты, которые в результате могут оказаться бракованными
- ☐ Прогноз спроса на наличные в АТМ банка: для каждого АТМ требуется спрогнозировать сколько наличных будет запрошено пользователями на фиксированном горизонте в будущем
- ☐ Определить жанр фильма по текстовой аннотации
- ☐ Анализ эмоциональной окраски сообщения в службе поддержки: определить носит ли сообщение позитивный, негативный или нейтральный характер

Задача ранжирования отличается от задачи регрессии следующим (укажите все правильные утверждения):

1 балл

- ☐ в задаче ранжирования относительный порядок числовых ответов важнее, чем абсолютное значение числового ответа, потому что ответ будет использоваться для упорядочивания объектов друг относительно друга
- ☐ в задаче ранжирования числовые ответы должны быть исключительно в интервале от 0 до 1
- ☐ в задаче ранжирования запрещено иметь одинаковые ответы на разных объектах, а в задаче регрессии это разрешено
- ☐ если умножить все ответы модели на 10, то в случае задачи ранжирования качество решения не ухудшится, а в случае задачи регрессии ухудшится существенно
- ☐ если умножить все ответы модели на 10, то в случае задачи регрессии качество решения не ухудшится, а в случае задачи ранжирования ухудшится существенно



При построении модели и оценке её качества, данные разделяют на обучающую, валидационную и тестовую выборку для того, чтобы (укажите все правильные утверждения) 1 балл

- ☐ Усложнить процесс построения модели. Достаточно деления данных на обучение и тест, выделение валидационной выборки никаких существенных плюсов не дает.
- ☐ Ускорить процесс обучения модели за счет того, что для обучения используется только обучающая выборка и она меньше по объему, чем все исходные данные
- ☐ Более точно оценить качество работы модели за счет того, что будет сделана оценка качества не только на тех данных, которые были продемонстрированы модели в процессе обучения, но и на тех данных, которые являются новыми для модели.
- ☐ Данные разделяют по принципу: наиболее качественные наблюдения отбираются в обучающую выборку, нетипичные объекты отправляют в валидационную выборку, а шумные наблюдения и выбросы идут в тест. За счет такой "фильтрации" данных для обучения достигается рост качества модели, а также валидация по нетипичным объектам позволяет получить надежную модель
- ☐ Снизить вероятность слишком тонкой подстройки параметров модели под фиксированную выборку: параметры подбираются по валидации, но модель дополнительно оценивается по тесту.

При подборе гиперпараметров моделей машинного обучения требуется проведение кросс-валидации. Какие из перечисленных ниже стратегий кросс-валидации применимы для оптимизации гиперпараметров модели? 1 балл

- ☐ стратегия k-fold
- ☐ стратегия leave-one-out (loo)
- ☐ стратегия tk-fold
- ☐ стратегия stratified loo
- ☐ стратегия k random splits
- ☐ стратегия k-loo
- ☐ стратегия stratified tk-fold



Каким образом строится финальная модель после подбора параметров методом кросс-валидации?

1 балл

- ☐ Каждый набор параметров тестируется на своем фолде, оценивается качество. Значения параметров усредняются с весами, пропорциональными качеству на фолдах. Именно с этими значениями параметров обучается финальная модель.
- ☐ Каждый набор параметров тестируется на своем фолде. Для финальной модели выбираются параметры с фолда, показавшего лучший результат по метрике. Далее с этими параметрами строится финальная модель.
- ☐ Каждый набор параметров тестируется на k фолдах, после чего оценка качества для каждого набора параметров усредняется по фолдам. Для финальной модели выбираются параметры, показавшие лучшее среднее на фолдах. Далее с этими параметрами строится финальная модель.
- ☐ Каждый набор параметров тестируется на k фолдах, после чего оценка качества для каждого набора параметров усредняется по фолдам. Значения параметров усредняются с весами, пропорциональными качеству на фолдах. Именно с этими значениями параметров обучается финальная модель.
- ☐ Каждый набор параметров тестируется на своем фолде, оценивается качество. Значения параметров усредняются. Именно с этими значениями параметров обучается финальная модель.

[Назад](#)

[Далее](#)

Никогда не используйте формы Google для передачи паролей.

Компания Google не имеет никакого отношения к этому контенту. [Сообщение о нарушении](#) - [Условия использования](#) - [Политика конфиденциальности](#)

Google Формы

