



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. Задача активного обучения состоит в том, чтобы

- ☐ ☐ использовать неразмеченную часть выборки (отклики неизвестны) для донастройки модели без узнавания их реальных откликов
- ☐ ☐ активнее использовать существующие размеченные объекты, неоднократно возвращаясь к ним в процессе обучения
- ☒ ☒ интеллектуально выбирать объекты из неразмеченной части выборки (метки неизвестны) для их доразметки, потом дообучать модель на расширенной обучающей выборке (с известными метками)

Балл: 2.0

Комментарий к правильному ответу:

2. Проблема холодного старта (cold start) в рекомендательных системах заключается

- ☒ ☒ в сложности построения рекомендаций новым пользователям и товарам, о рейтингах которых еще не собрано подробной статистики
- ☐ ☐ в сложности оценивания перепараметризованной рекомендательной системы
- ☐ ☐ в устойчивом смещении оценок пользователей в зависимости от их личной шкалы оценивания
- ☐ ☐ в смещении оценок товаров, вызванных фальшивыми оценками производителей этих товаров либо их конкурентами

Балл: 2

Комментарий к правильному ответу:

3. Факторизационные машины, в отличие от линейных моделей прогнозирования, позволяют учесть влияние на отклик

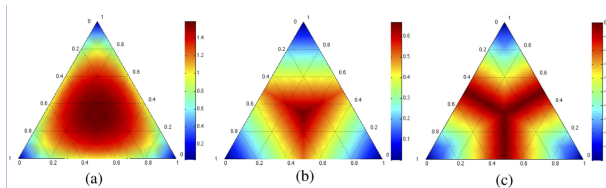
- ☐ ☐ тригонометрических преобразований от исходных признаков ($\sin(x)$, $\cos(x)$)
- ☐ ☐ любых нелинейных преобразований от исходных признаков
- ☐ ☐ квадратов исходных признаков
- ☒ ☒ попарных произведений различных признаков

Балл: 2

Комментарий к правильному ответу:



4. Пусть имеется задача классификации на 3 класса. Рассмотрим тепловые карты степени приоритетности объектов для разметки методами least confident sampling, margin sampling, entropy sampling. Какая из карт соответствует какому методу? Левый нижний угол - объект с вероятностями классов согласно модели $(1,0,0)$, верхний - $(0,1,0)$, правый нижний - $(0,0,1)$. Остальные точки соответствуют промежуточным вероятностям классов $(p(1), p(2), p(3))$ линейной интерполяцией.



- ☒ ☒ (a) - entropy sampling, (b) - least confident sampling, (c) - margin sampling
- ☐ ☐ (a) - least confident sampling, (b) - margin sampling, (c) - entropy sampling
- ☐ ☐ (a) - margin sampling, (b) - least confident sampling, (c) - entropy sampling
- ☐ ☐ (a) - least confident sampling, (b) - entropy sampling, (c) - margin sampling
- ☐ ☐ (a) - entropy sampling, (b) - margin sampling, (c) - least confident sampling
- ☐ ☐ (a) - margin sampling, (b) - entropy sampling, (c) - least confident sampling

Балл: 2

Комментарий к правильному ответу:

У margin $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ так же приоритетен, как $(\frac{1}{2}, \frac{1}{2}, 0)$, $(0, \frac{1}{2}, \frac{1}{2})$ и т.д. У least confident - линии уровня $p(i)=\text{const}$ - отрезки напротив i -й вершины.

5. Рассмотрим две валидационных выборки: в первой объекты сэмплируются случайно из генеральной совокупности, а во второй - выбираются методом entropy sampling активного обучения. При этом модель обучается на обучающей выборке, полученной случайным сэмплированием из генеральной совокупности и не дообучается на валидационных выборках - на них лишь тестируется качество модели. Как будут в общем случае связаны между собой точности модели на валидационных выборках, полученных каждым из способов?



- ☐ ☐ на второй выборке ожидается более высокая точность
- ☐ ☐ ожидается, что точности будут примерно соответствовать друг другу
- ☒ ☒ на второй выборке ожидается более низкая точность

Балл: 2

Комментарий к правильному ответу:

6. Построение рекомендаций с помощью матричных разложений в общем случае обладает преимуществом, по сравнению с рекомендациями посредством обычного сокращенного сингулярного разложения (pure SVD) в том, что при настройке модели

- ☒ ☐ используются только реально предоставленные рейтинги, а не все комбинации пользователь-товар в матрице рейтингов
- ☐ ☐ не учитываются пользователи и товары, у которых число рейтингов ниже заданного положительного порога
- ☐ ☒ появляется возможность аппроксимировать поведение пользователей матрицей более высокого ранга по сравнению с обычным сингулярным разложением

Балл: 0

Комментарий к правильному ответу:

7. Выберите метод активного обучения, напрямую применимый к регрессии:

- ☐ ☐ least confident sampling
- ☒ ☒ expected model change
- ☐ ☐ margin sampling
- ☐ ☐ entropy sampling

Балл: 2

Комментарий к правильному ответу:

8. Мера serendipity в рекомендательных системах показывает способность рекомендательной системы

- ☐ ☐ давать хорошие рекомендации для новых пользователей, еще не использовавших систему
- ☐ ☐ предоставлять разнообразные рекомендации, а не вариации одного и того же товара



- ☐ ☐ давать хорошие рекомендации для новых товаров, еще не рекомендовавшихся системой
- ☐ ☐ давать рекомендации, на которые пользователь будет часто кликать
- ☐ ☐ скрывать сомнительные рекомендации, которые пользователю могут не понравиться
- ☒ ☒ удивлять пользователя новыми нестандартными рекомендациями

Балл: 2

Комментарий к правильному ответу: