



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. При информационном поиске сформирована выдача из 5 документов, из 2 релевантных документов оба попали в выдачу на позиции 2 и 5. Вычислите меру $AP@5$.

Ответ: 0.45

Правильный ответ: 0.45

Погрешность: 0.0

Балл: 2.0

Комментарий к правильному ответу:

$$\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{4} + \frac{1}{5} = 0.25 + 0.2 = 0.45$$

2. Рассмотрим меру Discounted Cumulative Gain (DCG) с функцией расчета выигрыша $Gain(y) = 2^y - 1$ (y -степень релевантности) и функцией дисконтирования за более позднюю выдачу $Discount(i) = 1/i$, где $i=1,2,3,\dots$ - номер позиции. Вычислите нормализованную меру $nDCG@7$, если релевантными оказались 2й и 5й документ с $y=3$ и $y=4$ соответственно. Ответ введите с точностью до 2 знаков после запятой

Ответ: 0.35

Правильный ответ: 0.35

Погрешность: 0.001

Балл: 2.0

Комментарий к правильному ответу:

3. Почему на практике не используется минимизация доли неверно упорядоченных объектов напрямую при обучении попарных (pairwise) методов градиентными методами оптимизации?

- ☐ ☐ для доли дефектных пар, в отличие от др. функций потерь, не очевиден выбор начального приближения параметров модели
- ☒ ☒ этот функционал является кусочно-постоянным, что не позволяет его минимизировать градиентными методами
- ☐ ☐ используемые на практике функции потерь вычисляются на порядок быстрее для пар объектов, составляющих минибатч

Балл: 2.0

Комментарий к правильному ответу:



4. В рекомендательных системах для двух пар пользователей, корреляции между рейтингами которых равны в рамках пар, следует считать более похожей ту пару, у которых число товаров, одновременно оцененных обоими пользователями пары,

☐ ☐ это число не должно влиять на похожесть

☒ ☒ больше

☐ ☐ меньше

Балл: 2.0

Комментарий к правильному ответу:

5. В коллаборативной фильтрации, в отличие от рекомендательной системы общего вида, помимо данных о том, какие пользователи пользователи каким товарам выставили какие рейтинги

☒ ☒ нет информации о характеристиках пользователей и товаров, только об их историческом взаимодействии

☐ ☐ нет информации о характеристиках пользователей, но есть информация о характеристиках товаров

☐ ☐ нет информации о характеристиках товаров, но есть информация о характеристиках пользователей

Балл: 2.0

Комментарий к правильному ответу:

6. Неявные оценки товаров (implicit feedback) пользователями рекомендательной системы включают в себя:

☒ ☒ факт добавления товара в корзину

☒ ☒ факт покупки товара

☒ ☒ факт просмотра описания товара

☐ ☐ выставление оценки товару

☐ ☐ развернутый отзыв на товар

☐ ☐ выставление лайка/дизлайка товару

Балл: 2.0**Комментарий к правильному ответу:**

7. Рассмотрим меру Discounted Cumulative Gain (DCG) с функцией расчета выигрыша $\text{Gain}(y) = 2^y - 1$ (y -степень релевантности) и функцией дисконтирования за более позднюю выдачу $\text{Discount}(i) = 1/i$, где $i=1,2,3,\dots$ - номер позиции. Вычислите DCG@7, если релевантными оказались 2й и 5й документ с $y=3$ и $y=4$ соответственно.

Ответ: 6.5**Правильный ответ: 6.5****Погрешность: 0.0****Балл: 2.0****Комментарий к правильному ответу:**

8. При расширении выдачи на запрос пользователя в рекомендательной системе, в отличие от информационного поиска (information retrieval), делается предположение о том, что

☒ ☒ пользователь до конца не знает, что он ищет☐ ☐ пользователь чётко знает, что он ищет**Балл: 2.0****Комментарий к правильному ответу:**