



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 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:

2. Рассмотрим задачу ранжирования. Пусть `labeled` - обучающие выборки, где каждому объекту x поставлена в соответствие вещественная мера релевантности, а `preference` - обучающие выборки, где для подмножества пар объектов $(x(i), x(j))$ указана информация, какой из объектов пары более релевантен (но неизвестно насколько). Какие классы методов могут обучаться на каких типах обучающих выборок?

☐ ☐ поточечные (pointwise) методы - на `preference` и на `labeled`, попарные (pairwise) - только на `labeled`

☒ ☒ поточечные (pointwise) методы - только на `labeled`, попарные (pairwise) - на `preference` и на `labeled`

☐ ☐ поточечные (pointwise) методы - только на `preference`, попарные (pairwise) - только на `labeled`

☐ ☐ каждый метод может обучаться и на `labeled`, и на `preference`.

☐ ☐ поточечные (pointwise) методы - только на `labeled`, попарные (pairwise) - только на `preference`

Балл: 2.0

Комментарий к правильному ответу:



3. 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 нормализованную меру $n\text{DCG}@7$, 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:

4. В задаче ранжирования важность корректного порядка неравномерна - гораздо важнее правильно упорядочить объекты в начале выдачи (наиболее релевантные по версии модели), чем в конце. Выберите меры качества, учитывающие этот эффект, где K =полное число объектов в выдаче:

☒ ☒ AveragePrecision@K

☐ ☐ Precision@K

☒ ☒ MeanAveragePrecision@K

☒ ☒ nDCG@K

☐ ☐ доля дефектных (неверно упорядоченных) пар объектов при выдаче K объектов

☒ ☒ DCG@K

Балл: 2.0

Комментарий к правильному ответу:

5. Явные оценки товаров (explicit feedback) пользователями рекомендательной системы включают в себя:

☐ ☐ факт добавления товара в корзину

☒ ☒ развернутый отзыв на товар

☒ ☒ выставление оценки товару



☒ ☒ выставление лайка/дизлайка товару

☐ ☐ факт покупки товара

☐ ☐ факт просмотра описания товара

Балл: 2.0

Комментарий к правильному ответу:

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

☒ ☒ факт просмотра описания товара

☒ ☒ факт покупки товара

☐ ☐ выставление оценки товару

☒ ☒ факт добавления товара в корзину

☐ ☐ выставление лайка/дизлайка товару

☐ ☐ развернутый отзыв на товар

Балл: 2.0

Комментарий к правильному ответу:

7. 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 товаров, но есть информация о характеристиках пользователей

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

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

Балл: 2.0

Комментарий к правильному ответу:

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



- ☒ ☒ 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
- ☐ ☐ для доли дефектных пар, в отличие от др. функций потерь, не очевиден выбор начального приближения параметров модели

Балл: 2.0

Комментарий к правильному ответу: