



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. 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 - только на preference, попарные (pairwise) - только на labeled
- ☒ ☒ поточечные (pointwise) методы - только на labeled, попарные (pairwise) - на preference и на labeled
- ☐ ☐ поточечные (pointwise) методы - на preference и на labeled, попарные (pairwise) - только на labeled
- ☐ ☐ поточечные (pointwise) методы - только на labeled, попарные (pairwise) - только на preference
- ☐ ☐ каждый метод может обучаться и на labeled, и на preference.

**Балл: 2.0**

**Комментарий к правильному ответу:**

2. 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,

- ☒ ☒ больше
- ☐ ☐ это число не должно влиять на похожесть
- ☐ ☐ меньше

**Балл: 2.0**

**Комментарий к правильному ответу:**

3. 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:

4. 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  $\text{nDCG}@7$ , 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:

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

- ☐ ☐ факт просмотра описания товара
- ☐ ☐ факт покупки товара
- ☒ ☒ развернутый отзыв на товар
- ☒ ☒ выставление лайка/дизлайка товару
- ☐ ☐ факт добавления товара в корзину
- ☒ ☒ выставление оценки товару

Балл: 2.0

Комментарий к правильному ответу:

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

- ☒ ☒ факт добавления товара в корзину
- ☐ ☐ выставление оценки товару
- ☒ ☒ факт покупки товара
- ☒ ☒ факт просмотра описания товара
- ☐ ☐ развернутый отзыв на товар
- ☐ ☐ выставление лайка/дизлайка товару

Балл: 2.0



Комментарий к правильному ответу:

7. Рассмотрим линейную pairwise модель ранжирования  $a(x) = w^T x$ , обучаемую методом стохастического градиентного спуска с шагом обучения  $\eta > 0$ , сэмплированием одной пары объектов  $(i, j)$ ,  $i \succ j$  и экспоненциальной функцией потерь  $e^{-(a(x_i) - a(x_j))}$ . Выберите корректную формулу для обновления весов:

☐ ☐  $w := w + \eta e^{w^T(x_i - x_j)}(x_j - x_i)$

☐ ☐  $w := w + \eta e^{w^T(x_i - x_j)}(x_i - x_j)$

☐ ☒  $w := w + \eta e^{w^T(x_j - x_i)}(x_j - x_i)$

☒ ☐  $w := w + \eta e^{w^T(x_j - x_i)}(x_i - x_j)$

Балл: 0

Комментарий к правильному ответу:

8. 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:

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