

<p>Анализ неструктурированных данных Самостоятельная работа Максимальная оценка: 10 баллов</p>
--------------------------------------------------------------------------------------------------------

Имя и фамилия: \_\_\_\_\_

Группа: \_\_\_\_\_

**Ответьте на вопросы ниже. Можно использовать компьютер для вычислений и для просмотра лекций и учебников. Каждый вопрос оценивается в 2 балла.**

1. Можно ли напрямую оптимизировать метрики качества ранжирования? Почему?

Ответ: нельзя, т.к. метрики негладкие.

2. За счет чего достигается разнообразие поисковой выдачи?

Учитываем ...

1. разные смыслы слова-запроса
2. разные типы информации
3. жанровую специфику, в т.ч. “желтизну” контента
4. количество кликов пользователей на разные документы, которые находятся по запросу

Ответ: спорно, но 1, 2 и 4.

3. Могут ли векторные представления быть использованы для сопоставления объектов, различных по природе? Почему?

Ответ: можно. Аргументация лектора неизвестна, принимался любой разумный положительный ответ.

4. В какой постановке задачи ранжирования целевая функция определена на парах документов?

1. Pairwise-постановка задачи
2. Pointwise-постановка задачи
3. Listwise-постановка задачи

Ответ: 1

5. К какой постановке задачи ранжирования относится обучение ранжированию (learning to rank)?

1. Pairwise-постановка задачи
2. Pointwise-постановка задачи
3. Listwise-постановка задачи

Ответ: 3