

Отчет по домашнему заданию 5

Андрей Лебедев, группа 424

Задача 1

Эксперт нашёл 20 релевантных документов. Система нашла 4 документа в следующей последовательности релевантных и нерелевантных документов:

RNRNRRRNNNN

Решение

Для вычисления средней точности нужно определить долю релевантных документов на каждом этапе поиска и затем усреднить её.

$$AP = \frac{\frac{1}{3} \left(\frac{1}{1} + \frac{2}{3} + \frac{3}{6} + \frac{4}{7} \right)}{20} = \frac{\frac{1}{3} (1 + 0.67 + 0.5 + 0.57)}{20} = \frac{2.74}{20} = 0.137$$

Ответ: Средняя точность поиска составляет 0.137.

Задача 2

При разметке релевантных документов эксперты использовали шкалу от 0 до 4 баллов. При тестировании систем выяснилось, что системы выдали следующие результаты поиска ответов на один и тот же запрос:

- Система 1: 4, 2, 3, 1, 2, 0, ...
- Система 2: 3, 2, 4, 4, 4, 0, ...

Требуется определить, какая система ищет лучше по мере NDCG.

Решение

Для вычисления метрики $nDCG$ сначала найдём DCG для каждой системы по формуле:

$$DCG = \sum_{i=1}^n \frac{rel_i}{\log_2(i+1)}$$

Система 1:

$$\begin{aligned} DCG_1 &= \frac{4}{\log_2(1+1)} + \frac{2}{\log_2(2+1)} + \frac{3}{\log_2(3+1)} + \frac{1}{\log_2(4+1)} + \frac{2}{\log_2(5+1)} = \\ &= 4 + \frac{2}{1.585} + \frac{3}{2} + \frac{1}{2.322} + \frac{2}{2.585} \approx 4 + 1.26 + 1.5 + 0.43 + 0.77 = 7.96 \end{aligned}$$

Система 2:

$$\begin{aligned} DCG_2 &= \frac{3}{\log_2(1+1)} + \frac{2}{\log_2(2+1)} + \frac{4}{\log_2(3+1)} + \frac{4}{\log_2(4+1)} + \frac{4}{\log_2(5+1)} = \\ &= 3 + \frac{2}{1.585} + \frac{4}{2} + \frac{4}{2.322} + \frac{4}{2.585} \approx 3 + 1.26 + 2 + 1.72 + 1.55 = 9.53 \end{aligned}$$

Для нахождения $IDCG$ нам необходимо определить идеальное ранжирование документов. Поскольку точные документы, найденные первой и второй системами, неизвестны, обозначим $IDCG$ через некоторую константу C , которая будет одинаковой для обеих систем.

Тогда $nDCG$ для систем можно записать как:

$$nDCG_1 = \frac{7.96}{C}$$

$$nDCG_2 = \frac{9.53}{C}$$

Так как числитель у второй системы больше, получаем, что $nDCG_2 > nDCG_1$, следовательно, вторая система показывает лучшие результаты.

Ответ: По метрике $nDCG$ система 2 показывает лучший результат по сравнению с системой 1.