

Задание 1.

- 20 документов.
- RNRNNRRNNNN
- $AP = \frac{1+2/3+3/6+4/7}{20} = \frac{23}{168}$

Задание 2.

1. 4, 2, 3, 1, 2, 0 (и далее 0)
2. 3, 2, 4, 4, 4 (и далее 0)

- $\frac{NDCG1}{NDCG2} = \frac{DCG1}{DCG2}$
- $DCG1 = \frac{4}{\log_2(2)} + \frac{2}{\log_2(3)} + \frac{3}{\log_2(4)} + \frac{1}{\log_2(5)} + \frac{2}{\log_2(6)} \approx 7.9662$
- $DCG2 = \frac{3}{\log_2(2)} + \frac{2}{\log_2(3)} + \frac{4}{\log_2(4)} + \frac{4}{\log_2(5)} + \frac{4}{\log_2(6)} \approx 9.5319$
- $\frac{NDCG1}{NDCG2} = 0.8357$
- Система 2 лучше по NDCG.

Задание 3.

N = 1 000 000

Alpha = 0.7

Beta = 0.3

Терм	Df	Idf	Q	TF1	TF2	TF_IDF1	TF_IDF2
Отбор	70000	3.84	1	0	1	0	0.52
Кандидат	70000	3.84	1	1	0	0.5	0
Претендент	30000	5.06	0	1	1	0.66	0.68

Отобрать	50000	4.32	0	1	0	0.56	0
Выбрать	70000	3.84	0	0	1	0	0.52

$$q_m = \alpha * q + \frac{\beta}{2} (TFIDF1 + TFIDF2) = [0.78, 0.77, 0.20, 0.08, 0.08]$$