

Московский Авиационный Институт
(Научный Исследовательский Институт)

Факультет прикладной математики и физики
Кафедра вычислительной математики и программирования

Лабораторная работа №2 по курсу
«Информационный поиск»

Студент: Мхитарян С.А.

Преподаватель: Кухтичев А.А.

Группа: М8О-206М

Дата:

Оценка:

Подпись:

Москва, 2021

ЛР2: Оценка качества поиска

Задание

Необходимо оценить качество двух поисков и сравнить их с друг другом (для Википедии можно собственный поиск по Википедии, поиск Google или Яндекса с ограничением по сайту Википедии). Как минимум, нужно измерить P, DCG, NDCG и ERR уровней @1, @3 и @5, приветствуется использование дополнительных метрик качества. Для оценки качества необходимо придумать 30 запросов, отражающих интересы пользователей или, если есть доступ к настоящим запросам пользователей, то выбрать репрезентативную подборку.

Метод решения

1. Выбрать поисковики для сравнения друг с другом (выбраны поисковики Yandex и Wikipedia).
2. Придумать нетривиальные запросы
3. Провести оценку SERP для Яндекса с ограничением по википедии и для поиска в самой википедии.
4. По формулам вычислить качество результата двух данных поисковиков, посчитать средние значения и сделать вывод о лучшем поиске

Результаты выполнения

№	Текст запроса	Оценка SERP	
		Яндекс	Википедия
1	Армяне	5,5,5,5,5	5,5,5,5,1
2	Победитель the international	5,5,1,5,4	1,1,5,1,1
3	Самый старый памятник	1,1,4,3,4	3,1,1,4,4
4	Сколько костей у человека	5,1,1,5,3	1,1,1,5,1
5	Самая долгоживущая страна	5,4,2,5,1	1,1,1,1,1
6	Охота на ведьм	5,5,3,1,4	5,4,4,3,2
7	Самое высокое здание	5,5,5,3,2	5,4,5,1,5
8	Что такое ксенофобия	5,4,3,1,5	5,4,3,1,1
9	Преступление и наказание	5,5,5,4,3	5,5,4,3,3
10	Международная космическая станция площадь	5,5,4,2,5	4,3,1,1,1
11	Фаст фуд	5,5,5,5,5	5,1,2,2,1
12	Индоевропейская группа языков	5,2,1,5,5	5,3,2,2,2
13	день программиста	5,5,5,3,1	5,1,2,3,3
14	Гражданско-правовой кодекс	5,1,1,5,1	4,5,1,1,1
15	Когда правил Иван Грозный	5,5,4,1,3	5,4,5,5,5
16	На каком материке есть кенгуру	3,5,3,2,4	1,5,1,1,3
17	Первая жизнь на земле	5,1,1,4,1	1,1,1,1,5
18	Крым 2014	5,5,4,5,5	5,5,5,4,5
19	Когда умерли динозавры	5,3,5,3,5	4,1,3,1,1
20	Лысые кошки	5,1,5,4,5	1,5,5,1,1

21	Вакцина коронавирус	5,4,5,5,3	5,4,4,5,4
22	Сортировка пузырьком	5,5,5,5,5	5,3,1,5,1
23	Распад СССР	5,2,2,4,5	5,1,1,1,1
24	День благодарения	5,5,4,5,4	5,5,1,4,4
25	Где появилась первая карта	3,5,2,5,1	1,1,1,1,1
26	Буддизм	5,4,4,5,4	5,4,5,4,3
27	Адронный коллайдер	5,5,3,2,2	3,3,3,3,3
28	Москва сити	5,5,5,4,2	5,5,5,4,3
29	Испанка грипп	5,5,4,3,4	5,5,2,3,1
30	Самый большой самолёт в мире	5,1,4,1,1	5,4,2,1,4

Оценки поисковиков

№	Метрика	Яндекс			Википедия												
		@1	@3	@5	@1	@3	@5										
1	P	1.0 0	1.0 0	1.0 0	1.0 0	1.0 0	0.8 0	4	P	1.0 0	0.3 3	0.6 0	0.0 0	0.0 0	0.2 0	5	6
	CG	5.0 0	15. 0	25. 0	5.0 0	15. 0	24. 0		CG	5.0 0	7.0 0	15. 0	1.0 0	3.0 0	9.0 0		
	DCG	31. 0	55. 5	80. 9	31. 0	55. 5	69. 3		DCG	31. 0	32. 1	48. 2	1.0 0	2.1 3	15. 9		
	NDC G	6.2 0	3.7 0	3.2 4	6.2 0	3.7 0	2.7 8		NDC G	6.2 0	2.1 4	1.9 3	1.0 0	0.7 1	0.6 4		
	ERR	0.9 7	0.9 7	0.9 7	0.9 7	0.9 7	0.9 7		ERR	0.9 7	0.9 8	1.0 3	0.5 0	0.5 8	0.7 4		
2	1.00	1.0 0	1.0 0	1.0 0	0.6 7	0.8 0	1.0 0	5	P	1.0 0	0.6 7	0.8 0	0.0 0	0.0 0	0.0 0	6	7
	5.00	14. 0	23. 0	5.0 0	11. 0	19. 0	5.0 0		CG	5.0 0	11. 0	17. 0	1.0 0	3.0 0	5.0 0		
	31.0	51. 1	70. 3	31. 0	51. 1	63. 4	31. 0		DCG	31. 0	41. 0	54. 7	1.0 0	2.1 3	2.9 5		
	6.20	3.4 1	3.5 2	6.2 0	3.4 1	2.5 4	6.2 0		NDC G	6.2 0	2.7 3	2.1 6	1.0 0	0.7 1	0.5 9		
	0.97	0.9 8	0.9 8	0.9 7	0.9 9	1.0 1	0.9 7		ERR	0.9 7	0.9 8	0.9 8	0.5 0	0.5 8	0.5 8		
3	P	0.0 0	0,6 0	0.8 5	0.0 0	0.3 3	0.6 0	6	P	1.0 0	0.6 7	0.8 0	1.0 0	1.0 0	0.8 0	7	8
	CG	1.0 0	6.0 0	13. 0	3.0 0	5.0 0	13. 0		CG	5.0 0	13. 0	18. 0	5.0 0	13. 0	18. 0		
	DCG	1.0 0	9.1 3	18. 0	7.0 0	8.1 3	20, 4		DCG	31. 0	54. 1	60. 4	31. 0	58. 1	64. 4		
	NDC G	1.0 0	0.7 6	0.9 0	2.3 3	0.5 4	1.0 2		NDC G	6.2 0	3.6 1	2.4 2	6.2 0	3.8 7	2.5 8		
	ERR	0.5 0	1.0 7	1.0 2	0.8 8	0.5 0	0.3 3		ERR	0.9 7	0.9 9	0.9 9	0.9 7	0.9 8	0.9 8		
									P	1.0 0	0.6 7	0.8 0	1.0 0	1.0 0	0.8 0		

	CG	5.0 0	13. 0	18. 0	5.0 0	13. 0	18. 0
	DCG	31. 0	54. 1	60. 4	31. 0	58. 1	64. 4
	NDC G	6.2 0	3.6 1	2.4 2	6.2 0	3.8 7	2.5 8
	ERR	0.9 7	0.9 9	0.9 9	0.9 7	0.9 8	0.9 8
8	P	1.0 0	1.0 0	0.8 0	1.0 0	1.0 0	0.6 0
	CG	5.0 0	12. 0	18. 0	5.0 0	12. 0	14. 0
	DCG	31. 0	44. 0	56. 4	31. 0	44. 0	44. 8
	NDC G	6.2 0	2.9 3	2.2 6	6.2 0	2.9 3	1.7 9
	ERR	0.9 7	0.9 8	0.9 8	0.9 7	0.9 8	0.9 9
9	P	1.0 0	0.3 3	0.6 0	1.0 0	1.0 0	1.0 0
	CG	5.0 0	7.0 0	15. 0	5.0 0	14. 0	20. 0
	DCG	31. 0	50. 9	60. 1	31. 0	41. 0	94. 7
	NDC G	6.2 0	3.3 9	2.4 0	6.2 0	2.7 3	3.7 9
	ERR	0.9 7	0.9 8	0.9 8	0.9 7	0.9 8	1.0 7
10	P	1.0 0	1.0 0	0.8 0	1.0 0	1.0 0	0.8 0
	CG	5.0 0	14. 0	21. 0	3.0 0	6.0 0	8.0 0
	DCG	31. 0	58. 1	71. 4	7.0 0	9.4 0	10. 2
	NDC G	6.2 0	3.8 7	2.8 6	2.3 3	1.0 4	0.6 8
	ERR	0.9 7	1.0 0	1.0 2	0.8 8	0.9 0	0.9 4
11	P	1.0 0	1.0 0	1.0 0	1.0 0	0.3 3	0.2 0
	CG	5.0 0	15. 0	25. 0	5.0 0	8.0 0	11. 0
	DCG	31. 0	55. 5	80. 9	31. 0	33. 1	34. 8
	NDC G	6.2 0	3.7 0	3.2 4	6.2 0	2.2 1	1.3 9
	ERR	0.9 7	0.9 7	0.9 7	0.9 7	0.9 8	0.9 8
	P	1.0 0	1.0 0	1.0 0	1.0 0	0.3 3	0.2 0
	CG	5.0 0	15. 0	25. 0	5.0 0	8.0 0	11. 0
12	DCG	31. 0	55. 5	80. 9	31. 0	33. 1	34. 8
	NDC G	6.2 0	3.7 0	3.2 4	6.2 0	2.2 1	1.3 9
	ERR	0.9 7	0.9 7	0.9 7	0.9 7	0.9 8	0.9 8
	P	1.0 0	1.0 0	0.2 0	0.0 0	0.3 3	0.6 0
13	CG	5.0 0	15. 0	19. 0	2.0 0	8.0 0	14. 0
	DCG	31. 0	55. 5	58. 9	3.0 0	19. 1	24. 9
	NDC G	6.2 0	3.7 0	2.3 6	1.5 0	1.2 7	1.0 0
	ERR	0.9 7	0.9 7	0.9 6	0.7 5	0.9 8	0.9 9
14	P	1.0 0	0.3 3	0.6 0	1.0 0	0.6 7	0.4 0
	CG	5.0 0	7.0 0	13. 0	4.0 0	10. 0	12. 0
	DCG	31. 0	32. 1	46. 1	15. 0	35. 2	35. 9
	NDC G	6.2 0	2.1 4	1.8 4	3.7 5	2.3 5	1.4 4
15	ERR	0.9 7	0.9 8	0.9 8	0.9 4	0.9 7	0.9 7
	P	1.0 0	1.0 0	0.8 0	1.0 0	1.0 0	1.0 0
	CG	5.0 0	14. 0	18. 0	5.0 0	14. 0	24. 0
	DCG	31. 0	56. 0	59. 1	31. 0	56. 0	81. 4
	NDC G	6.2 0	3.7 3	2.3 6	6.2 0	3.7 3	3.2 6
	ERR	0.9 7	0.9 8	0.9 9	0.9 7	0.9 8	0.9 8

16	P	1.0 0	1.0 0	0.8 0	0.0 0	0.3 3	0.4 0		31.0	47. 1	65. 6	1.0 0	36. 1	36. 9	31. 0
	CG	3.0 0	11. 0	17. 0	1.0 0	5.0 0	11. 0		6.20	3.1 4	2.6 2	1.0 0	2.4 1	1.4 8	6.2 0
	DCG	7.0 0	30. 1	37. 2	1.0 0	5.9 3	18. 4		0.97	0.9 9	0.9 9	0.5 0	0.7 7	0.7 9	0.9 7
	NDC G	2.3 3	2.0 1	1.4 9	0.3 3	0.6 6	1.2 3		P	1.0 0	1.0 0	0.8 0	1.0 0	1.0 0	1.0 0
	ERR	0.8 8	0.8 9	0.8 9	0.5 0	0.7 7	0.7 9		CG	5.0 0	14. 0	22. 0	5.0 0	13. 0	22. 0
17	P	1.0 0	0.3 3	0.4 0	0.0 0	0.0 0	0.2 0	21	DCG	5.0 0	56. 0	72. 0	31. 0	41. 0	57. 0
	CG	5.0 0	7.0 0	12. 0	1.0 0	3.0 0	9.0 0		NDC G	1.0 0	3.7 3	2.8 8	6.2 0	2.7 3	2.2 8
	DCG	31. 0	32. 1	39. 0	1.0 0	2.1 3	14. 6		ERR	0.9 7	0.9 9	0.9 9	0.9 7	0.9 8	0.9 8
	NDC G	6.2 0	2.1 4	1.5 6	1.0 0	0.7 1	0.5 8		P	1.0 0	1.0 0	1.0 0	1.0 0	0.6 7	0.6 0
	ERR	0.9 7	0.9 8	0.9 8	0.5 0	0.5 8	0.6 0		CG	5.0 0	15. 0	25. 0	5.0 0	8.0 0	15. 0
18	P	1.0 0	1.0 0	1.0 0	1.0 0	1.0 0	1.0 0	22	DCG	31. 0	55. 5	80. 9	31. 0	35. 9	49. 7
	CG	5.0 0	14. 0	24. 0	4.0 0	14. 0	24. 0		NDC G	6.2 0	3.7 0	3.2 4	2.0 7	2.3 9	1.9 9
	DCG	31. 0	58. 1	83. 5	15. 0	50. 1	75. 5		ERR	0.9 7	0.9 7	0.9 7	0.9 7	0.9 7	0.9 7
	NDC G	6.2 0	3.8 7	3.3 4	3.7 5	3.3 4	3.0 2		P	1.0 0	0.3 3	0.6 0	0.0 0	0.3 3	0.2 0
	ERR	0.9 7	0.9 8	1.1 6	0.9 4	0.9 7	0.9 7		CG	4.0 0	8.0 0	17. 0	2.0 0	8.0 0	10. 0
19	P	1.0 0	1.0 0	1.0 0	1.0 0	0.6 7	0.4 0	23	DCG	15. 0	16. 1	34. 6	3.0 0	23. 1	23. 4
	CG	5.0 0	13. 0	21. 0	3.0 0	9.0 0	11. 0		NDC G	3.7 5	1.3 4	1.3 5	1.5 0	1.5 4	0.9 4
	DCG	31. 0	50. 9	66. 0	7.0 0	23. 1	24. 0		ERR	0.9 4	0.9 7	0.9 8	0.7 5	0.8 3	0.8 3
	NDC G	6.2 0	3.3 9	2.6 4	2.3 3	1.5 4	0.9 6		P	1.0 0	1.0 0	1.0 0	1.0 0	0.6 7	0.8 0
	ERR	0.9 7	0.9 8	0.9 9	0.8 8	0.9 1	1.0 6		CG	5.0 0	14. 0	23. 0	5.0 0	11. 0	19. 0
20	1.00	0.4 0	0.8 0	0.0 0	0.4 0	1.0 0	1.0 0	24	DCG	31. 0	51. 1	70. 3	31. 0	51. 1	63. 4
	5.00	11. 0	20. 0	1.0 0	11. 0	13. 0	5.0 0		NDC G	6.2 0	3.4 1	3.5 2	6.2 0	3.4 1	2.5 4

	ERR	0.9 7	0.9 8	0.9 8	0.9 7	0.9 9	1.0 1		CG	5.0 0	14. 0	21. 0	5.0 0	12. 0	16. 0
25	P	1.0 0	0.6 7	0.6 0	0.0 0	0.0 0	0.0 0		DCG	5.0 0	9.5 3	18. 4	31. 0	52. 1	55. 5
	CG	3.0 0	10. 0	16. 0	1.0 0	3.0 0	5.0 0		NDC G	1.0 0	0.6 4	0.7 4	6.2 0	3.4 7	2.2 4
	DCG	7.0 0	28. 1	41. 9	1.0 0	2.1 3	2.9 5		ERR	0.9 7	0.9 8	0.9 8	0.9 7	0.9 9	0.9 9
	NDC G	2.3 3	1.8 7	1.6 8	1.0 0	0.1 4	0.5 9		P	1.0 0	0.6 7	0.4 0	0.0 0	0.6 7	0.4 0
	ERR	0.8 8	1.1 1	1.1 3	0.5 0	0.5 8	0.5 8	30	CG	5.0 0	10. 0	12. 0	5.0 0	11. 0	16. 0
									DCG	31. 0	39. 1	40. 0	31. 0	42. 0	48. 2
26	P	1.0 0	1.0 0	1.0 0	1.0 0	1.0 0	0.8 0		NDC G	6.2 0	2.6 1	1.6 0	6.2 0	3.5 0	2.4 1
	CG	5.0 0	13. 0	22. 0	5.0 0	14. 0	21. 0		ERR	0.9 7	0.9 9	0.9 9	0.5 0	0.7 7	0.7 9
	DCG	31. 0	41. 0	57. 0	31. 0	56. 0	65. 2								
	NDC G	6.2 0	2.7 3	2.2 8	6.2 0	3.7 3	2.6 1								
	ERR	0.9 7	0.9 8	1.0 8	0.9 7	0.9 9	1.0 0								
27	P	0.0 0	0.3 3	0.6 0	1.0 0	1.0 0	1.0 0								
	CG	1.0 0	9.0 0	13. 0	3.0 0	9.0 0	15. 0								
	DCG	1.0 0	24. 1	26. 6	7.0 0	14. 9	20. 7								
	NDC G	1.0 0	1.6 1	1.0 6	2.3 3	1.6 6	1.3 8								
	ERR	0.5 0	0.7 4	0.7 5	0.8 8	1.1 8	1.2 0								
28	P	1.0 0	1.0 0	0.8 0	1.0 0	1.0 0	1.0 0								
	CG	5.0 0	15. 0	25. 0	5.0 0	15. 0	22. 0								
	DCG	31. 0	66. 1	91. 5	31. 0	66. 1	75. 3								
	NDC G	6.2 0	4.4 1	3.6 6	6.2 0	4.4 1	3.0 1								
	ERR	0.9 7	0.9 7	1.0 9	0.9 7	0.9 7	1.1 4								
29	P	1.0 0	1.0 0	1.0 0	1.0 0	0.6 7	0.8 0								

Средние значения

Метрика	Яндекс			Википедия		
	@1	@3	@5	@1	@3	@5
P	0.96	0.90	0.83	0.55	0.53	0.50
CG	3.9	8.56	17.61	4.21	10.11	16.65
DCG	44	41.02	51.28	15.3	29.13	39.87
NDCG	4.55	3.11	2.55	4.27	3.01	2.59
ERR	0.9	0.85	0.98	0.87	0.79	0.7

По таблице средних значений можно сделать вывод, что поиск Яндекса по Википедии работает лучше, чем нативный поиск Википедии. Поиск Яндекса лучше работает при нетривиальных запросах, в которых слова в разных падежах, имеются глаголы и предлоги. Но стоит отметить, что поиск Википедии работает лучше, если это название статьи или 1-2 слова в поиске.

В SERP и метрике P положительными считаются оценки > 2 .

В качестве дополнительной метрики была взята метрика CG.

Выводы

При выполнении лабораторной работы была проведена оценка качества работы двух поисковиков Яндекса и Википедии. Были посчитаны 5 метрик (P, CG, DCG, NDCG, ERR). Судя по полученным средним оценкам – поиск Яндекса работает на порядок лучше.