

Московский авиационный институт
Факультет прикладной математики и физики

Лабораторная работа №2

по курсу:
«Информационный поиск»
по теме:
«Оценка качества поиска»
2 семестр

Студент:	Ахмед С. Х.
Преподаватель:	Калинин А. Л.
Группа:	8О-106М

Москва, 2019 г

Постановка задачи

Необходимо оценить качество своего поиска и сравнить их с двумя альтернативами (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.

Провести анализ результатов оценки качества. Какие у какой поисковой системы сильные и слабые стороны? Как можно бороться с недостатками, что можно сделать, чтобы улучшить качество?

Оборудование:

Компьютер HP Omen 15 под управлением операционной системы Windows 10, Intel Core i5-7300HQ 2.50 GHz, 12 Gb RAM

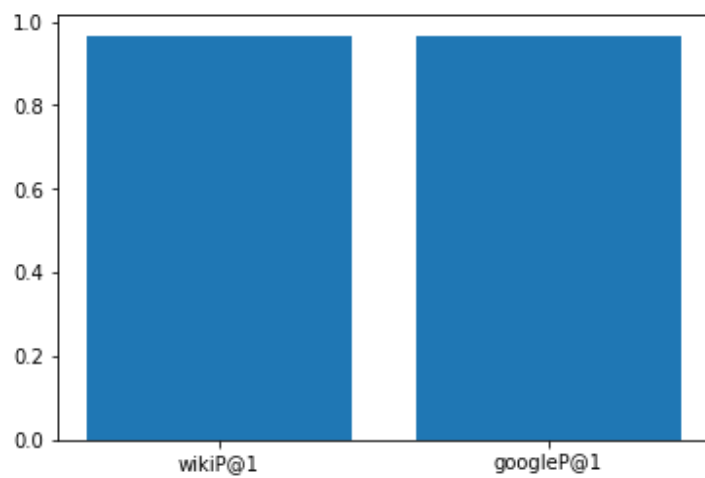
Программное обеспечение

Язык программирования	Python 3.6
Среда программирования	Anaconda, Jupyter Notebook

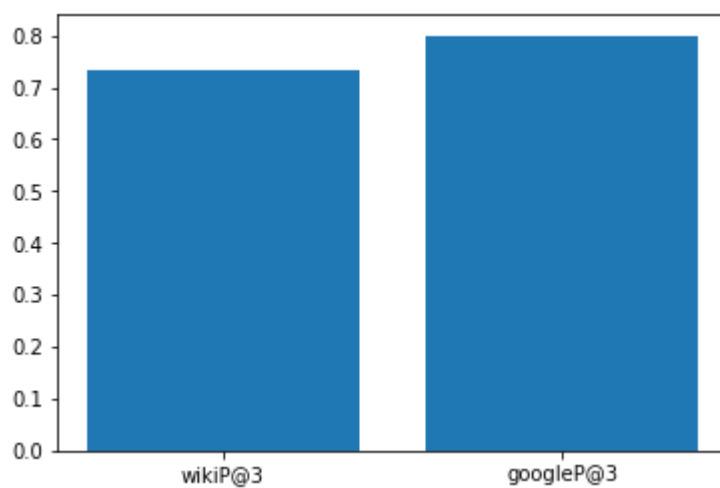
Полученные результаты

График Average Precision

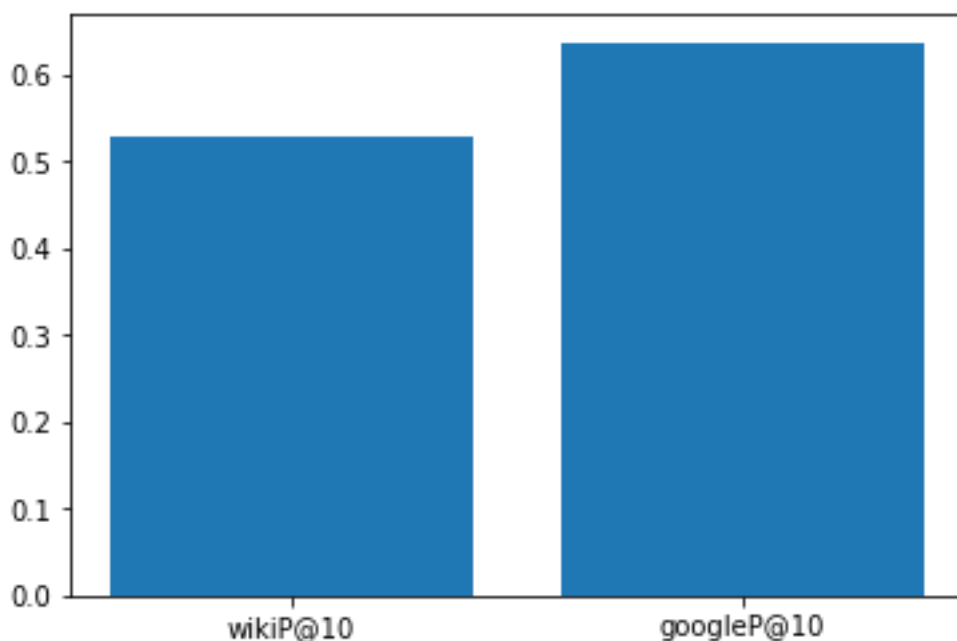
Код расположен в git. Здесь заметно, что поиск по Google (что неудивительно) имеет высокие средние показатели точности, (что неудивительно) в отличии от поиска Wikipedia. Здесь заметно, что отличие в качестве проявляется в метриках P@3 и P10. Обусловлено такое поведение тем, что у google больше информации относительно этих запросов (CTR, длительность сеанса, сведения на основе геолокации), чем у Wiki, а также в самом алгоритме поиска и ранжирования



Сравнение P@1 для Google и Wiki



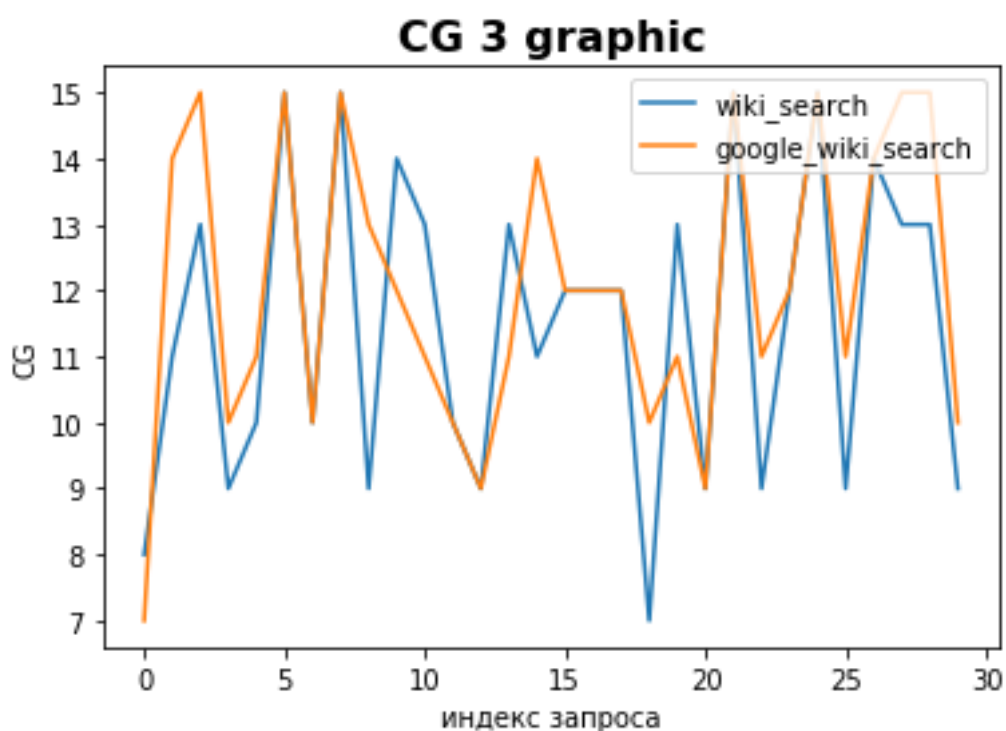
Сравнение P@3 для Google и Wiki

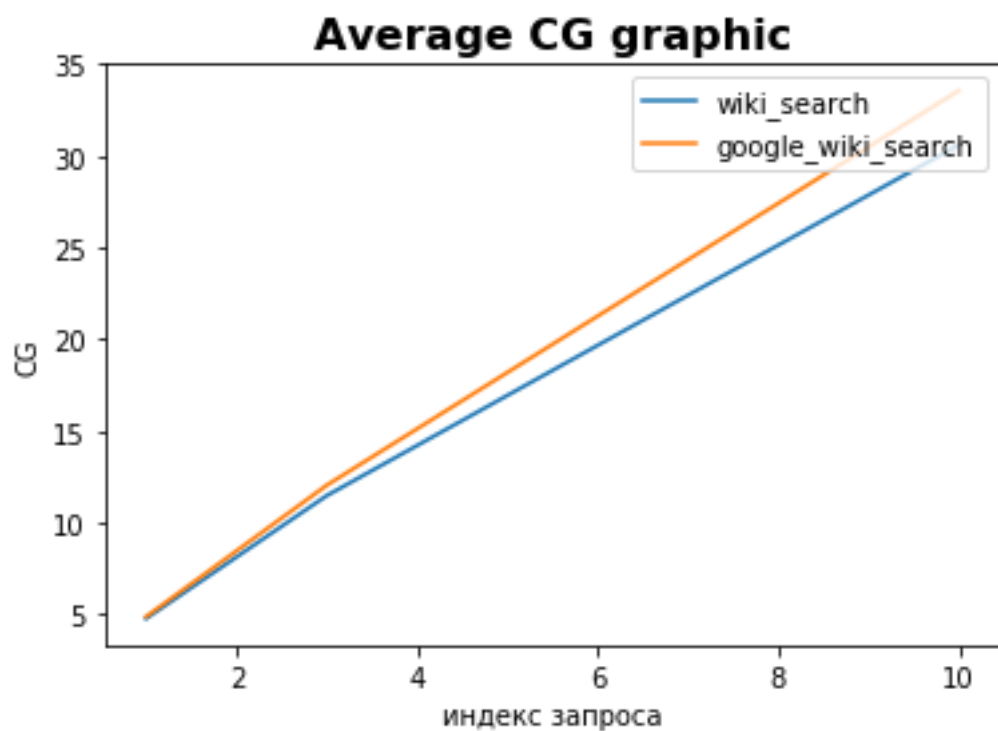
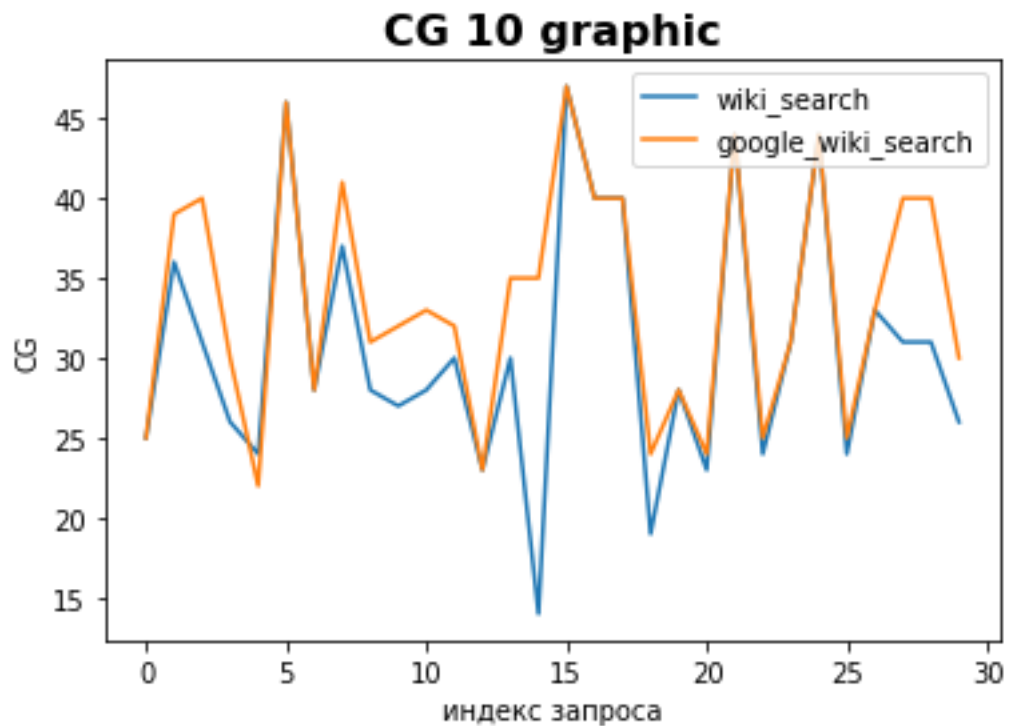


Сравнение P@10 для Google и Wiki

График CG@3 и GC@10

На данном изображении продемонстрировано распределение CG. Как можно заметить, разброс в полезности статей для Google. Оно очень отражает качество поискового вывода на этой выборке и неустойчивость в результатах. Когда усредним результаты, станет заметна общая тенденция: Полезность поисковой выдачи google стабильнее и выше. Особенно это заметно на CG@10. Картина, полученная при вычислении средних значений подтверждает это.

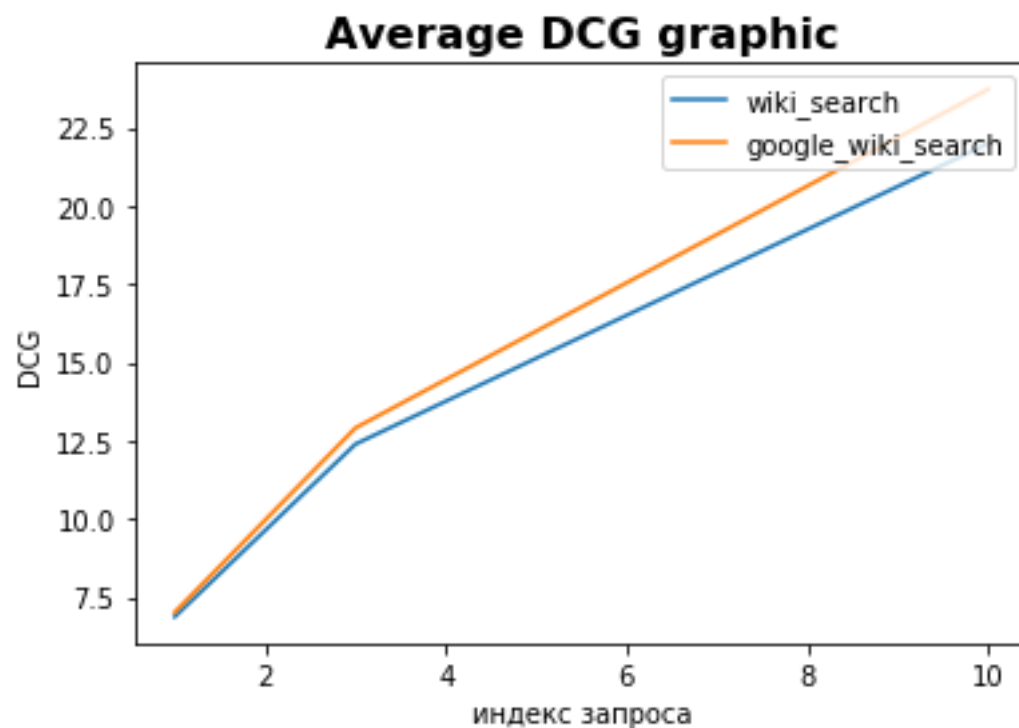
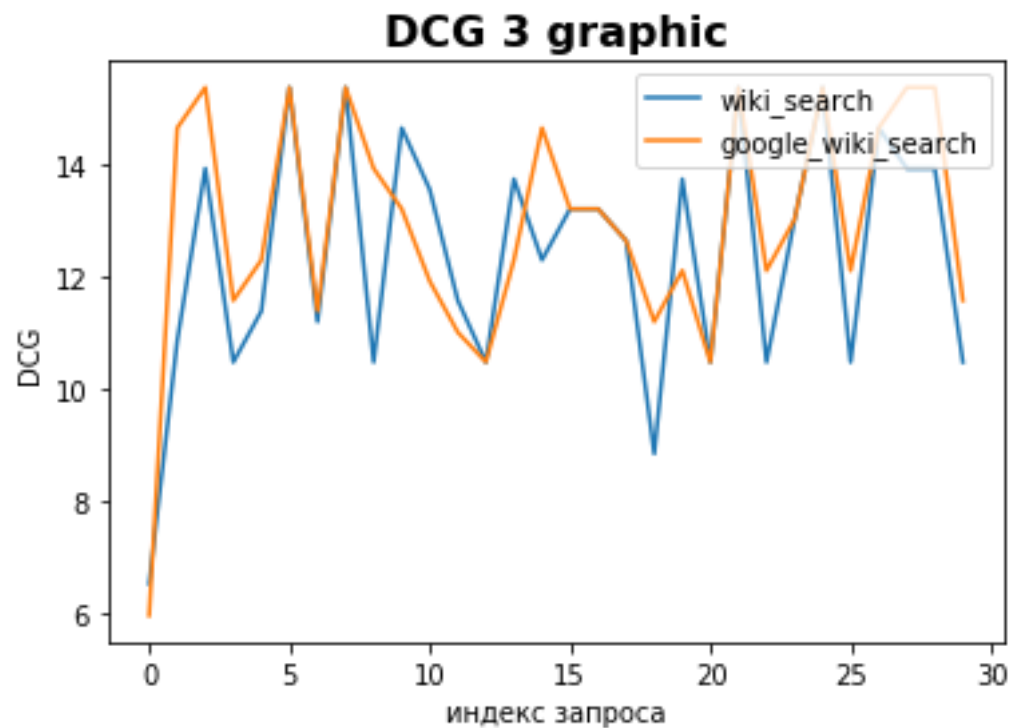




Аналогично с DCG

Сама оценка учитывает порядок релевантных документов, что является более качественной метрикой. В качестве коэффициента был взят логарифм по позиции. Судя по данным результатам стоит отметить следующее: Более явно заметно, что польза

поискового вывода google выше, то есть распределение релевантных запросов в google более равномерно



В NDCG учитывается наличие идеальной выдачи и происходит соотношение при той же величине: Здесь заметно в данном случае Google имеет выше оценку, что обусловлено, тем что идеальная выдача имеет большее соотношение оценки DCG текущей выдачи и идеальной

выдачи(отсортированной по оценкам)

