



گروه کامپیوتر - دانشکده مهندسی - دانشگاه فردوسی مشهد

مستندات پروژه اصلی

برای بخشی از فعالیتهای کلاسی

درس مبانی بازیابی اطلاعات و جست و جوی وب

موضوع:

بهبود سیستم ها پایه

TF-IDF و Language Model

ارائه دهنده:

سعید بهمنی سنگسری

افشین شهرستانی

استاد درس:

دکتر فائزه انسان

بهار 1398

فهرست مطالب:

3	مقدمه
4	روش ها
5	روش انتخابی
6	توضیح روش انتخابی
10	مقایسه معیارهای ارزیابی سیستم های مختلف
17	ارزیابی و جمع بندی

مقدمه:

در این سند ابتدا به مقایسه روش های موجود برای بهبود سیستم ها در lucene و سپس به راه حل انتخاب شده می پردازیم و در ادامه به کد نوشته شده در کتاب خانه lucene و مقایسه معیار های ارزیابی سیستم ها قبل و بعد از اعمال تغییر می پردازیم.

روش ها:

برای بهبود سیستم روش های متعددی موجود است که روش هایی که بررسی شد عبارت اند از: استفاده از BM25Similarity, position, page rank, domain authority.

:BM25Similarity

در اصل به این روش نسل بعدی TF-IDF می گویند که به معنی بهترین تطابق 25 که بسیار روش کار آمدی است ولی با توجه به مسئله اصلی (بهبود روش های پایه) استفاده از BM25Similarity مجاز نیست زیرا یک روش پیاده سازی شده است که در کتاب خانه lucene موجود و فقط نیاز به جایگزاری است و در اصل BM25Similarity یک سیستم مجزا از سیستم های پایه (TF-IDF, Language Model) است.

:Page rank

می توان هنگام استفاده از سیستم های پایه page rank را به فرمول نهایی اضافه کرد ولی باید به نرمال سازی توجه شود.

در page rank نیازمند لینک های ورودی و خروجی به هر صفحه را داشته و با استفاده از فرمول موجود می توان page rank را محاسبه کرد ولی با توجه به نبود لینک های ورودی به هر صفحه محاسبه page rank نیازمند به فرمول پیچیده و زمان بری بوده است.

:domain authority

در domain authority به کل domain وب سایت یک امتیاز اختصاص داده میشود و به فرمول نهایی اضافه کرد و باید به نرمال سازی توجه شود.

با توجه به این که دیتا تست شده بر روی سیستم از وب سایت Wikipedia است در نظر گرفتن domain authority فرقی در جواب سیستم بهبود یافته نمیکند، زیرا یک عدد ثابت به تمامی امتیاز ها اضافه میشود. پس در سیستم بهبود یافته که در نظر گرفته شده است از domain authority استفاده نشده است.

:Position

در position به ترتیب کلمات موجود در کوئری توجه میشود و برای ترتیب کلمات امتیازی در نظر گرفته میشود که در این مورد باید با چگونگی کار با stop word ها و نرمال سازی امتیاز توجه ویژه ای کرد.

روش انتخابی

با ذکر کردن ویژگی هریک از روش ها، روش انتخابی دخیل کردن position در سیستم های پایه شد. دلیل استفاده از این کار نزدیکی بسیار زیاد به صورت پروژه تعریف شده و قابل پیاده سازی بودن آن بوده است زیرا در BM25Similarity یک روش جدید است که ربطی به صورت پروژه نیست اگر در صورت پروژه فقط درخواست جواب بهتر بود در آن موقع استفاده از روش BM25Similarity مجاز و منطقی بود و برای بقیه موارد دلیل استفاده نشده در قبل ذکر شده است.

در روش انتخاب شده یک بار از روش قبلی جواب را پیدا کرده و سپس بوسیله position جواب دیگری بدست می آوریم و سپس امتیاز ها را بوسیله میانگین وزن دار مرتب کرده و 10 (در صورت پروژه برگرداندن 10 داکيومنت برتر بوده است) داکيومنت برتر را نمایش میدهیم. خوب بودن این روش به این صورت است که شاید کلمه کوئری که در داکيومنت برتر وجود داشته به تنهایی معنای جواب درست نبوده (در سیستم قبلی) و یا شاید کوئری شامل کلمه مهمی نبوده که فقط از position استفاده شود. با ترکیب کردن این دو روش هر یک نقص های دیگری را می پوشاند و جواب بهتری را نمایش میدهد.

توضیح روش انتخابی

برای روش های پایه هنگام صدا زدن `indexSearcher` به آن `parse` شده کلاس `QueryParser` را در آن وارد میگردیم و `ScoreDoc[]` در اینجا با توجه به ورودی `searcher`, 10 جواب برتر هستند

```
QueryParser qp = new QueryParser("contents", new StandardAnalyzer());
Query query = qp.parse(textToFind);
TopDocs standardHits = searcher.search(query, 10);
```

ولی برای دخیل کردن روش `position` از کلاس `PhraseQuery` می توان استفاده کرد ولی با توجه به تغییر هر نسخه کتاب خانه `lucene` کمی روش کار در هر نسخه فرق میکند. در `lucene8` برای استفاده از `PhraseQuery` از این نحو استفاده میکنیم.

```
PhraseQuery phraseQuery = new PhraseQuery(100,"contents",textToFind.split(" "));
TopDocs phraseHits = searcher.search(phraseQuery, 1000);
ScoreDoc[] phraseScoreDocs =phraseHits.scoreDocs;
```

در `lucene3.6` برای استفاده از `PhraseQuery` از این نحو استفاده میکنیم.

```
PhraseQuery phraseQuery = new PhraseQuery();
String[] strings=textToFind.split(" ");
for (int i=0;i<strings.length;i++) {
    phraseQuery.add(new Term(LuceneConstants.CONTENTES, strings[i]), i);
}
phraseQuery.setSlop(100);
TopDocs phraseHits = indexSearcher.search(phraseQuery, 10);
ScoreDoc[] phraseScoreDocs =phraseHits.scoreDocs;
```

که در اینجا متغیر `LuceneConstants.CONTENTES` برابر با مقدار رشته `contents` و در `setSlop` حداکثر فاصله هر `Term` است. که با توجه به اندازه هر داکيومنت و تحقیق درباره این موضوع مقداری بین 50 تا 100 معتبر است و زیر مقدار 50 نتایج محدودی تولید میکند. با توجه به تعداد داکيومنت های ایندکس شده که برابر 1000 داکيومنت است در مقدار دهی `setSlop` از مقدار 100 استفاده شده تا سیستم بهتر عمل کند. اگر تعداد داکيومنت ها بیش از 1000 بود مقدار `setSlop` برابر با عددی بین 10 تا 50 انتخاب میشد. اگر نیازمند پشت سر هم بودن و بدون فاصله کلمه های `query` بودیم از مقدار 0 در `setSlop` استفاده میشد که این کار باعث برگرداندن تعداد کم تری داکيومنت است.

در سیستم پیشنهادی یک بار با کوئری معمولی `TopDocs` را محاسبه کرده و بلافاصله بعد از آن با `phraseQuery` در یک متغیر جدید `TopDocs` را محاسبه میکنیم.

```
QueryParser qp = new QueryParser("contents", new StandardAnalyzer());
Query query = qp.parse(textToFind);
TopDocs standardHits = searcher.search(query, 100);
ScoreDoc[] standardScoreDoce = standardHits.scoreDocs;
PhraseQuery phraseQuery = new PhraseQuery(100, "contents", textToFind.split(""));
TopDocs phraseHits = searcher.search(phraseQuery, 100);
ScoreDoc[] phraseScoreDocs = phraseHits.scoreDocs;
```

در راه حلی که استفاده شده است مقدار ورودی دوم `indexSearcher` برابر 100 است زیرا اگر برابر با 10 میبود و ما هم 10 داکيومنت برتر را میخواستیم و با توجه با این که هر یک جواب مجزا تولید میکند (احتمال اشتراک داکيومنت ها وجود دارد) جواب بهینه ای برگردانده نمیشد.

```

Document standardDoc;

Document phraseDoc;

    boolean bool=false;

    for(int i=0;i<standardScoreDoce.length;i++){

        standardDoc= searcher.doc(standardScoreDoce[i].doc);

        ScoreDoc mergeDoc  =standardScoreDoce[i];

        for (int j=0;j<phraseScoreDocs.length;j++){

            phraseDoc= searcher.doc(phraseScoreDocs[j].doc);

            if(standardDoc.get("path").equals(phraseDoc.get("path"))){

                mergeDoc.score=(mergeDoc.score+phraseScoreDocs[j].score)/2;

                bool=true;

                break;

            }

        }

        if(!bool) {

            mergeDoc.score=mergeDoc.score/2;

        }

        list.add(mergeDoc);

        bool=false;

    }

```

بعد از مقداردهی 100 برای تعداد برگرداندن داکيومنت های برتر، سپس از حلقه تودرتو استفاده میکنیم که حلقه اول روی سیستم پایه و حلقه دوم روی سیستمی که از **position** استفاده میکند، این حلقه تودرتو داکيومنت مشترک در دو سیستم را پیدا کرده و میانگین امتیاز آن را محاسبه میکند که برای فهمیدن این که دو داکيومنت یکی هستند یا نه محل ذخیره شده آن را با هم مقایسه میکند .

و در **if** آخر اگر داکيومنتی پیدا نشد امتیاز داکيومنت سیستم پایه را تقسیم بر 2 میکند((امتیاز داکيومنت پایه + 0)/2 که 0 برای پیدا نشدن داکيومنت در سیستم **position** است) و در آخر داکيومنت را در آرایه اضافه میکند.


```
list.sort(new Comparator<ScoreDoc>() {
    @Override
    public int compare(ScoreDoc o1, ScoreDoc o2) {
        if(o1.score>o2.score) {
            return -1;
        }else if(o1.score<o2.score){
            return 1;
        }
        return 0;
    }
});
```

سپس آرایه ای که در آن داکيومنت ها موجود است را بر اساس امتیاز آن مرتب میکنیم چون با توجه به عدد 100 که در قبل درباره آن صحبت شد، در این آرایه 100 داکيومنت موجود است و ما 10 برتر را نیاز داریم، به همین دلیل آن را مرتب میکنیم.

```
List<ScoreDoc>finalList=new ArrayList<>();
for(int i=0;i<TopNDocs;i++){
    if(i<list.size()){
        finalList.add(list.get(i));
    }
}
```

بعد از این که در آرایه داکيومنت ها را بر اساس امتیاز مرتب کردیم، با توجه به مقدار TopNDocs که برابر تعداد برترین داکيومنت ها (در اینجا مقدار آن 10 است) برای نمایش است یک حلقه ایجاد میکنیم تا در آرایه جدید بریزیم.

```
return finalList;
```

در نهایت لیست نهایی که شامش TopNDocs عنصر است را برمیگردانیم.

مقایسه معیارهای ارزیابی سیستم های مختلف

در این قسمت خروجی سیستم پایه و سیستم جدید (ترکیب **position** با سیستم پایه) را مورد بررسی قرار

میدهیم.

ارزیابی بر روی سیستم پایه (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 0.0000	1	relative_unranked_avg_prec5	all 0.3600
num_ret	all 10	2	ircl_prn.0.80	all 0.0000	2	relative_unranked_avg_prec10	all 0.2500
num_rel	all 10	3	ircl_prn.0.90	all 0.0000	3	relative_unranked_avg_prec15	all 0.1111
num_rel_ret	all 5	4	ircl_prn.1.00	all 0.0000	4	relative_unranked_avg_prec20	all 0.0625
map	all 0.4196	5	P5	all 0.6000	5	relative_unranked_avg_prec30	all 0.0278
gm_ap	all 0.4196	6	P10	all 0.5000	6	relative_unranked_avg_prec100	all 0.0025
R-prec	all 0.5000	7	P15	all 0.3333	7	relative_unranked_avg_prec200	all 0.0006
bpref	all 0.5000	8	P20	all 0.2500	8	relative_unranked_avg_prec500	all 0.0001
recip_rank	all 1.0000	9	P30	all 0.1667	9	relative_unranked_avg_prec1000	all 0.0000
num_nonrel_judged_ret	all 0	10	P100	all 0.0500	10	utility 1.0 -1.0 0.0 0.0	all 0.0000
exact_prec	all 0.5000	11	P200	all 0.0250	11	rcl_at_142_nonrel	all 0.4958
exact_recall	all 0.5000	12	P500	all 0.0100	12	fallout_recall_0	all 0.3000
11-pt_avg	all 0.4773	13	P1000	all 0.0050	13	fallout_recall_14	all 0.5000
3-pt_avg	all 0.5417	14	recall15	all 0.3000	14	fallout_recall_28	all 0.5000
avg_doc_prec	all 0.4196	15	recall110	all 0.5000	15	fallout_recall_42	all 0.5000
exact_relative_prec	all 0.5000	16	recall115	all 0.5000	16	fallout_recall_56	all 0.5000
avg_relative_prec	all 0.7102	17	recall120	all 0.5000	17	fallout_recall_71	all 0.5000
exact_unranked_avg_prec	all 0.2500	18	recall130	all 0.5000	18	fallout_recall_85	all 0.5000
exact_relative_unranked_avg_prec	all 0.2500	19	recall1100	all 0.5000	19	fallout_recall_99	all 0.5000
map_at_R	all 0.7602	20	recall200	all 0.5000	20	fallout_recall_113	all 0.5000
int_map	all 0.4250	21	recall500	all 0.5000	21	fallout_recall_127	all 0.5000
exact_int_R_rcl_prec	all 0.5000	22	recall1000	all 0.5000	22	fallout_recall_142	all 0.5000
int_map_at_R	all 0.7806	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 0.5000	24	0.40R-prec	all 0.7500	24	int_0.40R-prec	all 0.7500
bpref_retnonrel	all 0.5000	25	0.60R-prec	all 0.5000	25	int_0.60R-prec	all 0.6250
bpref_topnonrel	all 0.5000	26	0.80R-prec	all 0.6250	26	int_0.80R-prec	all 0.6250
bpref_top5Rnonrel	all -nan	27	1.00R-prec	all 0.5000	27	int_1.00R-prec	all 0.5000
bpref_top10Rnonrel	all -nan	28	1.20R-prec	all 0.4167	28	int_1.20R-prec	all 0.4167
bpref_top10pRnonrel	all 0.5000	29	1.40R-prec	all 0.3571	29	int_1.40R-prec	all 0.3571
bpref_top25pRnonrel	all 0.5000	30	1.60R-prec	all 0.3125	30	int_1.60R-prec	all 0.3125
bpref_top50pRnonrel	all 0.5000	31	1.80R-prec	all 0.2778	31	int_1.80R-prec	all 0.2778
bpref_top25p2Rnonrel	all 0.5000	32	2.00R-prec	all 0.2500	32	int_2.00R-prec	all 0.2500
bpref_recall	all 1.0000	33	relative_prec5	all 0.6000	33	micro_prec	all 0.5000
bpref_5	all 1.0000	34	relative_prec10	all 0.5000	34	micro_recall	all 0.5000
bpref_10	all 1.0000	35	relative_prec15	all 0.5000	35	micro_bpref	all -nan
bpref_num_all	all 0.0000	36	relative_prec20	all 0.5000			
bpref_num_ret	all 0.0000	37	relative_prec30	all 0.5000			
bpref_num_correct	all 0	38	relative_prec100	all 0.5000			
bpref_num_possible	all 0	39	relative_prec200	all 0.5000			
old_bpref	all 0.5000	40	relative_prec500	all 0.5000			
old_bpref_top10pRnonrel	all 0.5000	41	relative_prec1000	all 0.5000			
infAP	all 0.4196	42	unranked_avg_prec5	all 0.1800			
gm_bpref	all 0.5000	43	unranked_avg_prec10	all 0.2500			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.1667			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.1250			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.0833			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.0250			
ircl_prn.0.40	all 0.6250	48	unranked_avg_prec200	all 0.0125			
ircl_prn.0.50	all 0.6250	49	unranked_avg_prec500	all 0.0050			
ircl_prn.0.60	all 0.0000	50	unranked_avg_prec1000	all 0.0025			
		51					

ارزیابی بر روی سیستم بهبود یافته (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 1.0000	1	relative_unranked_avg_prec5	all 1.0000
num_ret	all 10	2	ircl_prn.0.80	all 1.0000	2	relative_unranked_avg_prec10	all 1.0000
num_rel	all 10	3	ircl_prn.0.90	all 1.0000	3	relative_unranked_avg_prec15	all 0.4444
num_rel_ret	all 10	4	ircl_prn.1.00	all 1.0000	4	relative_unranked_avg_prec20	all 0.2500
map	all 1.0000	5	P5	all 1.0000	5	relative_unranked_avg_prec30	all 0.1111
gm_ap	all 1.0000	6	P10	all 1.0000	6	relative_unranked_avg_prec100	all 0.0100
R-prec	all 1.0000	7	P15	all 0.6667	7	relative_unranked_avg_prec200	all 0.0025
bpref	all 1.0000	8	P20	all 0.5000	8	relative_unranked_avg_prec500	all 0.0004
recip_rank	all 1.0000	9	P30	all 0.3333	9	relative_unranked_avg_prec1000	all 0.0001
num_nonrel_judged_ret	all 0	10	P100	all 0.1000	10	utility_1.0_-1.0_0.0_0.0	all 10.0000
exact_prec	all 1.0000	11	P200	all 0.0500	11	rcl_at_142_nonrel	all 1.0000
exact_recall	all 1.0000	12	P500	all 0.0200	12	fallout_recall_0	all 0.0000
11-pt_avg	all 1.0000	13	P1000	all 0.0100	13	fallout_recall_14	all 1.0000
3-pt_avg	all 1.0000	14	recall15	all 0.5000	14	fallout_recall_28	all 1.0000
avg_doc_prec	all 1.0000	15	recall10	all 1.0000	15	fallout_recall_42	all 1.0000
exact_relative_prec	all 1.0000	16	recall115	all 1.0000	16	fallout_recall_56	all 1.0000
avg_relative_prec	all 1.0000	17	recall120	all 1.0000	17	fallout_recall_71	all 1.0000
exact_unranked_avg_prec	all 1.0000	18	recall130	all 1.0000	18	fallout_recall_85	all 1.0000
exact_relative_unranked_avg_prec	all 1.0000	19	recall100	all 1.0000	19	fallout_recall_99	all 1.0000
map_at_R	all 1.0000	20	recall200	all 1.0000	20	fallout_recall_113	all 1.0000
int_map	all 1.0000	21	recall500	all 1.0000	21	fallout_recall_127	all 1.0000
exact_int_R_rcl_prec	all 1.0000	22	recall1000	all 1.0000	22	fallout_recall_142	all 1.0000
int_map_at_R	all 1.0000	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 1.0000	24	0.40R-prec	all 1.0000	24	int_0.40R-prec	all 1.0000
bpref_retnonrel	all 1.0000	25	0.60R-prec	all 1.0000	25	int_0.60R-prec	all 1.0000
bpref_topnonrel	all 1.0000	26	0.80R-prec	all 1.0000	26	int_0.80R-prec	all 1.0000
bpref_top5Rnonrel	all -nan	27	1.00R-prec	all 1.0000	27	int_1.00R-prec	all 1.0000
bpref_top10Rnonrel	all -nan	28	1.20R-prec	all 0.8333	28	int_1.20R-prec	all 0.8333
bpref_top10pRnonrel	all 1.0000	29	1.40R-prec	all 0.7143	29	int_1.40R-prec	all 0.7143
bpref_top25pRnonrel	all 1.0000	30	1.60R-prec	all 0.6250	30	int_1.60R-prec	all 0.6250
bpref_top50pRnonrel	all 1.0000	31	1.80R-prec	all 0.5556	31	int_1.80R-prec	all 0.5556
bpref_top25p2Rnonrel	all 1.0000	32	2.00R-prec	all 0.5000	32	int_2.00R-prec	all 0.5000
bpref_retail	all 1.0000	33	relative_prec5	all 1.0000	33	micro_prec	all 1.0000
bpref_5	all 1.0000	34	relative_prec10	all 1.0000	34	micro_recall	all 1.0000
bpref_10	all 1.0000	35	relative_prec15	all 1.0000	35	micro_bpref	all -nan
bpref_num_all	all 0.0000	36	relative_prec20	all 1.0000			
bpref_num_ret	all 0.0000	37	relative_prec30	all 1.0000			
bpref_num_correct	all 0	38	relative_prec100	all 1.0000			
bpref_num_possible	all 0	39	relative_prec200	all 1.0000			
old_bpref	all 1.0000	40	relative_prec500	all 1.0000			
old_bpref_top10pRnonrel	all 1.0000	41	relative_prec1000	all 1.0000			
infAP	all 1.0000	42	unranked_avg_prec5	all 0.5000			
gm_bpref	all 1.0000	43	unranked_avg_prec10	all 1.0000			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.6667			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.5000			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.3333			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.1000			
ircl_prn.0.40	all 1.0000	48	unranked_avg_prec200	all 0.0500			
ircl_prn.0.50	all 1.0000	49	unranked_avg_prec500	all 0.0200			
ircl_prn.0.60	all 1.0000	50	unranked_avg_prec1000	all 0.0100			
		51					

ارزیابی بر روی سیستم پایه (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 0.7000	1	relative_unranked_avg_prec5	all 1.0000
num_ret	all 10	2	ircl_prn.0.80	all 0.0000	2	relative_unranked_avg_prec10	all 0.6049
num_rel	all 9	3	ircl_prn.0.90	all 0.0000	3	relative_unranked_avg_prec15	all 0.2178
num_rel_ret	all 7	4	ircl_prn.1.00	all 0.0000	4	relative_unranked_avg_prec20	all 0.1225
map	all 0.7286	5	P5	all 1.0000	5	relative_unranked_avg_prec30	all 0.0544
gm_ap	all 0.7286	6	P10	all 0.7000	6	relative_unranked_avg_prec100	all 0.0049
R-prec	all 0.6667	7	P15	all 0.4667	7	relative_unranked_avg_prec200	all 0.0012
bpref	all 0.7778	8	P20	all 0.3500	8	relative_unranked_avg_prec500	all 0.0002
recip_rank	all 1.0000	9	P30	all 0.2333	9	relative_unranked_avg_prec1000	all 0.0000
num_nonrel_judged_ret	all 0	10	P100	all 0.0700	10	utility_1.0_-1.0_0.0_0.0	all 4.0000
exact_prec	all 0.7000	11	P200	all 0.0350	11	rcl_at_142_nonrel	all 0.7746
exact_recall	all 0.7778	12	P500	all 0.0140	12	fallout_recall_0	all 0.5556
11-pt_avg	all 0.6870	13	P1000	all 0.0070	13	fallout_recall_14	all 0.7778
3-pt_avg	all 0.6667	14	recall5	all 0.5556	14	fallout_recall_28	all 0.7778
avg_doc_prec	all 0.7286	15	recall10	all 0.7778	15	fallout_recall_42	all 0.7778
exact_relative_prec	all 0.7778	16	recall15	all 0.7778	16	fallout_recall_56	all 0.7778
avg_relative_prec	all 0.8885	17	recall20	all 0.7778	17	fallout_recall_71	all 0.7778
exact_unranked_avg_prec	all 0.5444	18	recall30	all 0.7778	18	fallout_recall_85	all 0.7778
exact_relative_unranked_avg_prec	all 0.4900	19	recall100	all 0.7778	19	fallout_recall_99	all 0.7778
map_at_R	all 0.9378	20	recall200	all 0.7778	20	fallout_recall_113	all 0.7778
int_map	all 0.7286	21	recall500	all 0.7778	21	fallout_recall_127	all 0.7778
exact_int_R_rcl_prec	all 0.7000	22	recall1000	all 0.7778	22	fallout_recall_142	all 0.7778
int_map_at_R	all 0.9405	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 0.7778	24	0.40R-prec	all 1.0000	24	int_0.40R-prec	all 1.0000
bpref_retnonrel	all 0.7778	25	0.60R-prec	all 0.8333	25	int_0.60R-prec	all 0.8571
bpref_topnonrel	all 0.7778	26	0.80R-prec	all 0.7500	26	int_0.80R-prec	all 0.7500
bpref_top5Rnonrel	all 0.7778	27	1.00R-prec	all 0.6667	27	int_1.00R-prec	all 0.7000
bpref_top10Rnonrel	all 0.7778	28	1.20R-prec	all 0.6364	28	int_1.20R-prec	all 0.6364
bpref_top10pRnonrel	all 0.7778	29	1.40R-prec	all 0.5385	29	int_1.40R-prec	all 0.5385
bpref_top25pRnonrel	all 0.7778	30	1.60R-prec	all 0.4667	30	int_1.60R-prec	all 0.4667
bpref_top50pRnonrel	all 0.7778	31	1.80R-prec	all 0.4118	31	int_1.80R-prec	all 0.4118
bpref_top25p2Rnonrel	all 0.7778	32	2.00R-prec	all 0.3889	32	int_2.00R-prec	all 0.3889
bpref_recall	all 1.0000	33	relative_prec5	all 1.0000	33	micro_prec	all 0.7000
bpref_5	all 1.0000	34	relative_prec10	all 0.7778	34	micro_recall	all 0.7778
bpref_10	all 1.0000	35	relative_prec15	all 0.7778	35	micro_bpref	all 0.0000
bpref_num_all	all 7.0000	36	relative_prec20	all 0.7778			
bpref_num_ret	all 0.0000	37	relative_prec30	all 0.7778			
bpref_num_correct	all 0	38	relative_prec100	all 0.7778			
bpref_num_possible	all 9	39	relative_prec200	all 0.7778			
old_bpref	all 0.7778	40	relative_prec500	all 0.7778			
old_bpref_top10pRnonrel	all 0.7778	41	relative_prec1000	all 0.7778			
infAP	all 0.7286	42	unranked_avg_prec5	all 0.5556			
gm_bpref	all 0.7778	43	unranked_avg_prec10	all 0.5444			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.3630			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.2722			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.1815			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.0544			
ircl_prn.0.40	all 1.0000	48	unranked_avg_prec200	all 0.0272			
ircl_prn.0.50	all 1.0000	49	unranked_avg_prec500	all 0.0109			
ircl_prn.0.60	all 0.8571	50	unranked_avg_prec1000	all 0.0054			
		51					

ارزیابی بر روی سیستم بهبود یافته (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 1.0000	1	relative_unranked_avg_prec5	all 1.0000
num_ret	all 10	2	ircl_prn.0.80	all 1.0000	2	relative_unranked_avg_prec10	all 1.0000
num_rel	all 10	3	ircl_prn.0.90	all 1.0000	3	relative_unranked_avg_prec15	all 0.4444
num_rel_ret	all 10	4	ircl_prn.1.00	all 1.0000	4	relative_unranked_avg_prec20	all 0.2500
map	all 1.0000	5	P5	all 1.0000	5	relative_unranked_avg_prec30	all 0.1111
gm_ap	all 1.0000	6	P10	all 1.0000	6	relative_unranked_avg_prec100	all 0.0100
R-prec	all 1.0000	7	P15	all 0.6667	7	relative_unranked_avg_prec200	all 0.0025
bpref	all 1.0000	8	P20	all 0.5000	8	relative_unranked_avg_prec500	all 0.0004
recip_rank	all 1.0000	9	P30	all 0.3333	9	relative_unranked_avg_prec1000	all 0.0001
num_nonrel_judged_ret	all 0	10	P100	all 0.1000	10	utility_1.0_-1.0_0.0_0.0	all 10.0000
exact_prec	all 1.0000	11	P200	all 0.0500	11	rcl_at_142_nonrel	all 1.0000
exact_recall	all 1.0000	12	P500	all 0.0200	12	fallout_recall_0	all 0.0000
11-pt_avg	all 1.0000	13	P1000	all 0.0100	13	fallout_recall_14	all 1.0000
3-pt_avg	all 1.0000	14	recall5	all 0.5000	14	fallout_recall_28	all 1.0000
avg_doc_prec	all 1.0000	15	recall10	all 1.0000	15	fallout_recall_42	all 1.0000
exact_relative_prec	all 1.0000	16	recall15	all 1.0000	16	fallout_recall_56	all 1.0000
avg_relative_prec	all 1.0000	17	recall20	all 1.0000	17	fallout_recall_71	all 1.0000
exact_unranked_avg_prec	all 1.0000	18	recall30	all 1.0000	18	fallout_recall_85	all 1.0000
exact_relative_unranked_avg_prec	all 1.0000	19	recall100	all 1.0000	19	fallout_recall_99	all 1.0000
map_at_R	all 1.0000	20	recall200	all 1.0000	20	fallout_recall_113	all 1.0000
int_map	all 1.0000	21	recall500	all 1.0000	21	fallout_recall_127	all 1.0000
exact_int_R_rcl_prec	all 1.0000	22	recall1000	all 1.0000	22	fallout_recall_142	all 1.0000
int_map_at_R	all 1.0000	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 1.0000	24	0.40R-prec	all 1.0000	24	int_0.40R-prec	all 1.0000
bpref_retnonrel	all 1.0000	25	0.60R-prec	all 1.0000	25	int_0.60R-prec	all 1.0000
bpref_topnonrel	all 1.0000	26	0.80R-prec	all 1.0000	26	int_0.80R-prec	all 1.0000
bpref_top5Rnonrel	all -nan	27	1.00R-prec	all 1.0000	27	int_1.00R-prec	all 1.0000
bpref_top10Rnonrel	all -nan	28	1.20R-prec	all 0.8333	28	int_1.20R-prec	all 0.8333
bpref_top10pRnonrel	all 1.0000	29	1.40R-prec	all 0.7143	29	int_1.40R-prec	all 0.7143
bpref_top25pRnonrel	all 1.0000	30	1.60R-prec	all 0.6250	30	int_1.60R-prec	all 0.6250
bpref_top50pRnonrel	all 1.0000	31	1.80R-prec	all 0.5556	31	int_1.80R-prec	all 0.5556
bpref_top25p2Rnonrel	all 1.0000	32	2.00R-prec	all 0.5000	32	int_2.00R-prec	all 0.5000
bpref_recall	all 1.0000	33	relative_prec5	all 1.0000	33	micro_prec	all 1.0000
bpref_5	all 1.0000	34	relative_prec10	all 1.0000	34	micro_recall	all 1.0000
bpref_10	all 1.0000	35	relative_prec15	all 1.0000	35	micro_bpref	all -nan
bpref_num_all	all 0.0000	36	relative_prec20	all 1.0000			
bpref_num_ret	all 0.0000	37	relative_prec30	all 1.0000			
bpref_num_correct	all 0	38	relative_prec100	all 1.0000			
bpref_num_possible	all 0	39	relative_prec200	all 1.0000			
old_bpref	all 1.0000	40	relative_prec500	all 1.0000			
old_bpref_top10pRnonrel	all 1.0000	41	relative_prec1000	all 1.0000			
infAP	all 1.0000	42	unranked_avg_prec5	all 0.5000			
gm_bpref	all 1.0000	43	unranked_avg_prec10	all 1.0000			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.6667			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.5000			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.3333			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.1000			
ircl_prn.0.40	all 1.0000	48	unranked_avg_prec200	all 0.0500			
ircl_prn.0.50	all 1.0000	49	unranked_avg_prec500	all 0.0200			
ircl_prn.0.60	all 1.0000	50	unranked_avg_prec1000	all 0.0100			

ارزیابی بر روی سیستم پایه (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 0.0000	1	relative_unranked_avg_prec5	all 0.6400
num_ret	all 10	2	ircl_prn.0.80	all 0.0000	2	relative_unranked_avg_prec10	all 0.3906
num_rel	all 8	3	ircl_prn.0.90	all 0.0000	3	relative_unranked_avg_prec15	all 0.1111
num_rel_ret	all 5	4	ircl_prn.1.00	all 0.0000	4	relative_unranked_avg_prec20	all 0.0625
map	all 0.5531	5	P5	all 0.8000	5	relative_unranked_avg_prec30	all 0.0278
gm_ap	all 0.5531	6	P10	all 0.5000	6	relative_unranked_avg_prec100	all 0.0025
R-prec	all 0.6250	7	P15	all 0.3333	7	relative_unranked_avg_prec200	all 0.0006
bpref	all 0.5000	8	P20	all 0.2500	8	relative_unranked_avg_prec500	all 0.0001
recip_rank	all 1.0000	9	P30	all 0.1667	9	relative_unranked_avg_prec1000	all 0.0000
num_nonrel_judged_ret	all 1	10	P100	all 0.0500	10	utility 1.0 -1.0 0.0 0.0	all 0.0000
exact_prec	all 0.5000	11	P200	all 0.0250	11	rcl_at_142_nonrel	all 0.6215
exact_recall	all 0.6250	12	P500	all 0.0100	12	fallout_recall_0	all 0.3750
11-pt_avg	all 0.5659	13	P1000	all 0.0050	13	fallout_recall_14	all 0.6250
3-pt_avg	all 0.6000	14	recall15	all 0.5000	14	fallout_recall_28	all 0.6250
avg_doc_prec	all 0.5531	15	recall10	all 0.6250	15	fallout_recall_42	all 0.6250
exact_relative_prec	all 0.6250	16	recall15	all 0.6250	16	fallout_recall_56	all 0.6250
avg_relative_prec	all 0.7663	17	recall20	all 0.6250	17	fallout_recall_71	all 0.6250
exact_unranked_avg_prec	all 0.3125	18	recall30	all 0.6250	18	fallout_recall_85	all 0.6250
exact_relative_unranked_avg_prec	all 0.2500	19	recall100	all 0.6250	19	fallout_recall_99	all 0.6250
map_at_R	all 0.8485	20	recall200	all 0.6250	20	fallout_recall_113	all 0.6250
int_map	all 0.5531	21	recall500	all 0.6250	21	fallout_recall_127	all 0.6250
exact_int_R_rcl_prec	all 0.6250	22	recall1000	all 0.6250	22	fallout_recall_142	all 0.6250
int_map_at_R	all 0.8615	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 0.5000	24	0.40R-prec	all 0.7500	24	int_0.40R-prec	all 0.8000
bpref_retnonrel	all 0.3750	25	0.60R-prec	all 0.8000	25	int_0.60R-prec	all 0.8000
bpref_topnonrel	all 0.5000	26	0.80R-prec	all 0.5714	26	int_0.80R-prec	all 0.6250
bpref_top5Rnonrel	all 0.5000	27	1.00R-prec	all 0.6250	27	int_1.00R-prec	all 0.6250
bpref_top10Rnonrel	all 0.5000	28	1.20R-prec	all 0.5000	28	int_1.20R-prec	all 0.5000
bpref_top10pRnonrel	all 0.5000	29	1.40R-prec	all 0.4167	29	int_1.40R-prec	all 0.4167
bpref_top25pRnonrel	all 0.5000	30	1.60R-prec	all 0.3846	30	int_1.60R-prec	all 0.3846
bpref_top50pRnonrel	all 0.5000	31	1.80R-prec	all 0.3333	31	int_1.80R-prec	all 0.3333
bpref_top25p2Rnonrel	all 0.5000	32	2.00R-prec	all 0.3125	32	int_2.00R-prec	all 0.3125
bpref_recall	all 0.6000	33	relative_prec5	all 0.8000	33	micro_prec	all 0.5000
bpref_5	all 0.8000	34	relative_prec10	all 0.6250	34	micro_recall	all 0.6250
bpref_10	all 0.8000	35	relative_prec15	all 0.6250	35	micro_bpref	all 0.1250
bpref_num_all	all 8.0000	36	relative_prec20	all 0.6250			
bpref_num_ret	all 3.0000	37	relative_prec30	all 0.6250			
bpref_num_correct	all 2	38	relative_prec100	all 0.6250			
bpref_num_possible	all 16	39	relative_prec200	all 0.6250			
old_bpref	all 0.3750	40	relative_prec500	all 0.6250			
old_bpref_top10pRnonrel	all 0.3750	41	relative_prec1000	all 0.6250			
infAP	all 0.5531	42	unranked_avg_prec5	all 0.4000			
gm_bpref	all 0.5000	43	unranked_avg_prec10	all 0.3125			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.2083			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.1562			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.1042			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.0312			
ircl_prn.0.40	all 0.8000	48	unranked_avg_prec200	all 0.0156			
ircl_prn.0.50	all 0.8000	49	unranked_avg_prec500	all 0.0062			
ircl_prn.0.60	all 0.6250	50	unranked_avg_prec1000	all 0.0031			

ارزیابی بر روی سیستم بهبود یافته (tf-idf)

num_q	all 1	1	ircl_prn.0.70	all 0.9000	1	relative_unranked_avg_prec5	all 1.0000
num_ret	all 10	2	ircl_prn.0.80	all 0.9000	2	relative_unranked_avg_prec10	all 1.0000
num_rel	all 9	3	ircl_prn.0.90	all 0.9000	3	relative_unranked_avg_prec15	all 0.3600
num_rel_ret	all 9	4	ircl_prn.1.00	all 0.9000	4	relative_unranked_avg_prec20	all 0.2025
map	all 0.9627	5	P5	all 1.0000	5	relative_unranked_avg_prec30	all 0.0900
gm_ap	all 0.9627	6	P10	all 0.9000	6	relative_unranked_avg_prec100	all 0.0081
R-prec	all 0.8889	7	P15	all 0.6000	7	relative_unranked_avg_prec200	all 0.0020
bpref	all 0.6667	8	P20	all 0.4500	8	relative_unranked_avg_prec500	all 0.0003
recip_rank	all 1.0000	9	P30	all 0.3000	9	relative_unranked_avg_prec1000	all 0.0001
num_nonrel_judged_ret	all 1	10	P100	all 0.0900	10	utility_1.0_-1.0_0.0_0.0	all 8.0000
exact_prec	all 0.9000	11	P200	all 0.0450	11	rcl_at_142_nonrel	all 0.9977
exact_recall	all 1.0000	12	P500	all 0.0180	12	fallout_recall_0	all 0.6667
11-pt_avg	all 0.9636	13	P1000	all 0.0090	13	fallout_recall_14	all 1.0000
3-pt_avg	all 0.9667	14	recall5	all 0.5556	14	fallout_recall_28	all 1.0000
avg_doc_prec	all 0.9627	15	recall10	all 1.0000	15	fallout_recall_42	all 1.0000
exact_relative_prec	all 1.0000	16	recall15	all 1.0000	16	fallout_recall_56	all 1.0000
avg_relative_prec	all 0.9621	17	recall20	all 1.0000	17	fallout_recall_71	all 1.0000
exact_unranked_avg_prec	all 0.9000	18	recall30	all 1.0000	18	fallout_recall_85	all 1.0000
exact_relative_unranked_avg_prec	all 0.8100	19	recall100	all 1.0000	19	fallout_recall_99	all 1.0000
map_at_R	all 0.9702	20	recall200	all 1.0000	20	fallout_recall_113	all 1.0000
int_map	all 0.9667	21	recall500	all 1.0000	21	fallout_recall_127	all 1.0000
exact_int_R_rcl_prec	all 0.9000	22	recall1000	all 1.0000	22	fallout_recall_142	all 1.0000
int_map_at_R	all 0.9778	23	0.20R-prec	all 1.0000	23	int_0.20R-prec	all 1.0000
bpref_allnonrel	all 0.6667	24	0.40R-prec	all 1.0000	24	int_0.40R-prec	all 1.0000
bpref_retnonrel	all 0.6667	25	0.60R-prec	all 1.0000	25	int_0.60R-prec	all 1.0000
bpref_topnonrel	all 0.6667	26	0.80R-prec	all 0.8750	26	int_0.80R-prec	all 0.9000
bpref_top5Rnonrel	all 0.6667	27	1.00R-prec	all 0.8889	27	int_1.00R-prec	all 0.9000
bpref_top10Rnonrel	all 0.6667	28	1.20R-prec	all 0.8182	28	int_1.20R-prec	all 0.8182
bpref_top10pRnonrel	all 0.6667	29	1.40R-prec	all 0.6923	29	int_1.40R-prec	all 0.6923
bpref_top25pRnonrel	all 0.6667	30	1.60R-prec	all 0.6000	30	int_1.60R-prec	all 0.6000
bpref_top50pRnonrel	all 0.6667	31	1.80R-prec	all 0.5294	31	int_1.80R-prec	all 0.5294
bpref_top25p2Rnonrel	all 0.6667	32	2.00R-prec	all 0.5000	32	int_2.00R-prec	all 0.5000
bpref_recall	all 0.6667	33	relative_prec5	all 1.0000	33	micro_prec	all 0.9000
bpref_5	all 1.0000	34	relative_prec10	all 1.0000	34	micro_recall	all 1.0000
bpref_10	all 0.6667	35	relative_prec15	all 1.0000	35	micro_bpref	all 0.0000
bpref_num_all	all 6.0000	36	relative_prec20	all 1.0000			
bpref_num_ret	all 6.0000	37	relative_prec30	all 1.0000			
bpref_num_correct	all 0	38	relative_prec100	all 1.0000			
bpref_num_possible	all 9	39	relative_prec200	all 1.0000			
old_bpref	all 0.6667	40	relative_prec500	all 1.0000			
old_bpref_top10pRnonrel	all 0.6667	41	relative_prec1000	all 1.0000			
infAP	all 0.9627	42	unranked_avg_prec5	all 0.5556			
gm_bpref	all 0.6667	43	unranked_avg_prec10	all 0.9000			
ircl_prn.0.00	all 1.0000	44	unranked_avg_prec15	all 0.6000			
ircl_prn.0.10	all 1.0000	45	unranked_avg_prec20	all 0.4500			
ircl_prn.0.20	all 1.0000	46	unranked_avg_prec30	all 0.3000			
ircl_prn.0.30	all 1.0000	47	unranked_avg_prec100	all 0.0900			
ircl_prn.0.40	all 1.0000	48	unranked_avg_prec200	all 0.0450			
ircl_prn.0.50	all 1.0000	49	unranked_avg_prec500	all 0.0180			
ircl_prn.0.60	all 1.0000	50	unranked_avg_prec1000	all 0.0090			
		51					

ارزیابی و جمع بندی

با توجه به نتایج سیستم پایه (tf-idf) و بهبود یافته، سیستم بهبود یافته نتایج بسیار بهتری را نشان میدهد که نشان از بهتر بودن سیستم جدید با مقایسه سیستم پایه است.

برای سیستم پایه (Language Model) در مقایسه با سیستم بهبود یافته، سیستم بهبود یافته کمی بهتر نشان میداد زیرا جواب های خود سیستم پایه خوب بود.

در آخر باید به این نکته اشاره کرد که اگر داکيومنت ها از وب سایت های مختلف جمع آوری شده بودند دخیل کردن domain authority در اکثر اوقات باعث بهبود جواب های برگردانده شده می شود.