OKSAT at NTCIR-14 OpenLiveQ-2

- Reorder Questions by Using White and Black -

Takashi SATO Yuki NAGASE Masato URAJI
sato@cc.osaka-kyoiku.ac.jp
(Osaka Kyoiku University)

[0] Outline

- Introduction
- White and Black Words
- Submitted Runs
- Evaluation
- Conclusions

[1] Introduction

- OKSAT submitted 18 runs for NTCIR-14 OpenLiveQ-2 task.
- We reorder questions by using white and black words because most questions in OpenliveQ2-question-data of this task are fit for the queries.
- The white words are selected by the frequency in questions, Google suggest and/or manual.
- On the other hand, the black words are selected by the rareness in the questions.
- The reorder of questions by white words is more effective than by black words from the evaluation results.

[2] White and Black Words

- The question data (OpenliveQ2-question-data.tsv) has provided almost matching questions for each query.
- As we have described in the runs of section 3, the Q-measure does not change much even if the top 10 questions for each query are moved to the bottom (run-U0) or all questions are sorted in the reverse order (run-U2).
- Therefore, we defined nouns that appear in questions and seem to fit to a query as white words. Conversely we defined nouns that seems difficult to fit to a query as black words.
- By moving the questions containing white words forward in the ranked list and moving the questions containing black words backward in the list, we considered that the questions fitted to the query can be gathered around top of the list.
- We set nouns which appear less frequently in the questions for each query as black words. On the other hand, we tried to set nouns as white words, (1) which appear many times in questions, (2) which are suggested by using Google suggestion, (3) which we found by manual, and (4) which are found in Wikipedia.

[2-1] White and Black Words Cnt'd

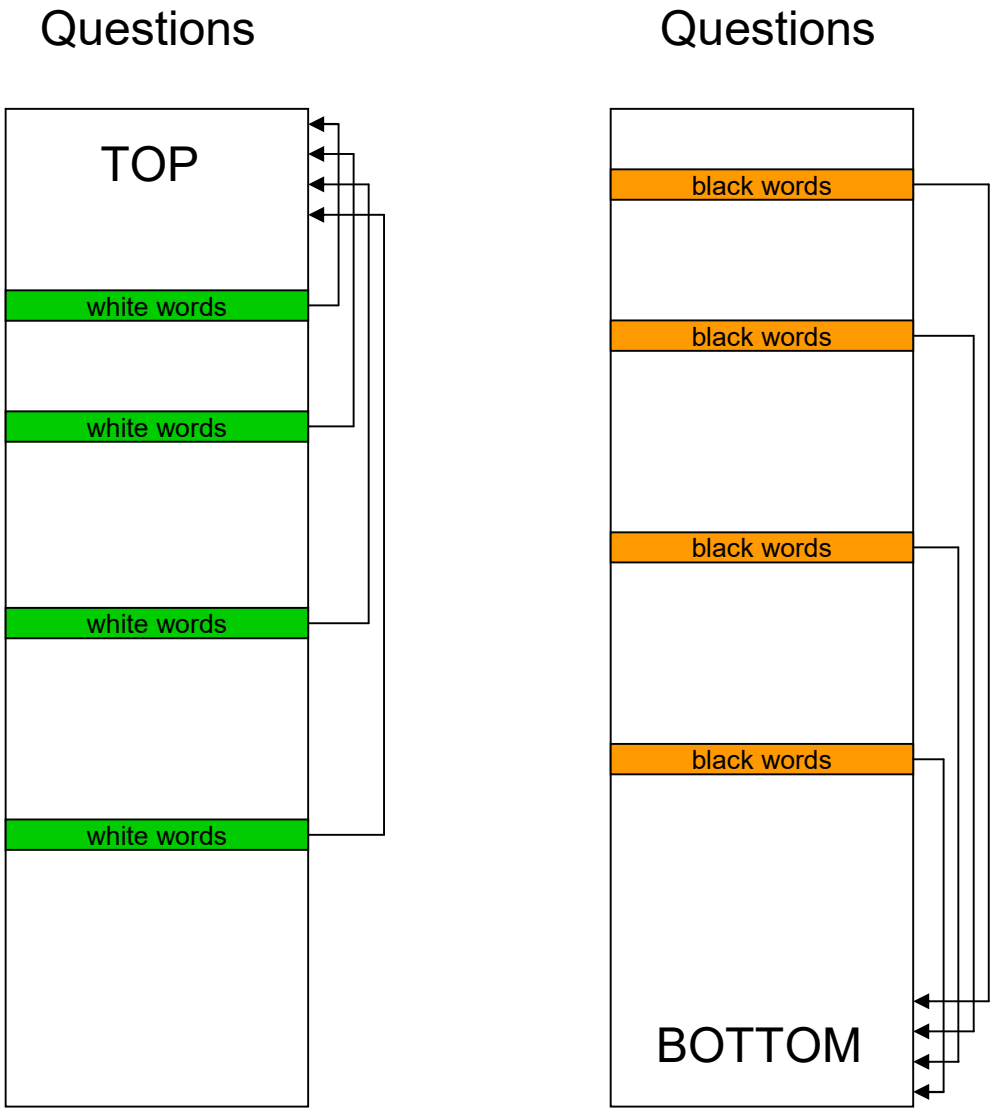

Fig. 1. Moving questions by white and black words.

[3] Submitted Runs

We submitted the following 18 runs. The numbers in parentheses are the formal run id's of the task.

- run-S0(90): The same run as ORG(89), i.e. the questions are ranked by the same order in the question-data.
- run-S1(108): We ranked the questions by using the probabilistic model[3] retrieving the title of questions by the query words.
- run-S2(121): Same ranking method were used in run-S1, however a parameter of probabilistic model was adjusted.

[3-1] Submitted Runs Cnt'd

- run-S3(132): The questions which have the plural black words (rare (≦3) noun words in the questions for each query) were ranked low in the ranked list. We extracted nouns by morphological analysis [4][5] of the title.
- run-S4(151): Same ranking method were used in run-S3, however we changed the standard of rareness (≦4).

[from OpenliveQ-1]

- run9(138): The same ranking method as run-9 of OKSAT [6] in OpenLiveQ-1. Questions were ranked by (page view)/(square root of body length).
- run20(135): The same ranking method as run-20 of OKSAT in OpenLiveQ-1.

[3-2] Submitted Runs Cnt'd

- run-N3(114): A white word based run, using the Google Suggest API to extract related terms and sort in descending order of the number of occurrences in the questions.
- run-N4(119): A black word based run, with questions including proper nouns appearing only once in all questions for the query, followed by the rest in the initial order.
- run-N5(142): Minor change version of run-N4. Run that dynamically adjusted the number of occurrences of black words so that more than half of the original TOP10 questions contain black words.
- run-N6(146): After separating the blacklist with run-N4 equivalent, the run that applied the sort of white word corresponding to run-N3 to non-blacklist.
- run-N7(153): Minor change version of run-N6. Run was evaluated by adding the white word including rate and the number of page view normalized by their maximum values.

[3-3] Submitted Runs Cnt'd

Table 1. White words by using Google Suggest API.

query ID	query	white words
OLQ-1001	博多	観光, ホテル, 映画, 天気, ラーメン, お土産, ランチ, グルメ, 屋台, もつ鍋
OLQ-1002	mt車	中古, 運転, レンタカー, 新車, メリット, おすすめ, 練習, コツ, クラッチ, 軽自動車
OLQ-1003	逃走中	ハンター, 動画, 2018, アプリ, やらせ, ゲーム, ヒカキン, の動画, ドラゴンボール, 上野
…	…	…

[3-4] Submitted Runs Cnt'd

- run-U0(94):We moved the top 10 questions behind the bottom.
- run-U1(96):We sorted questions in descending order by the number of page view in questions for each query.
- run-U2(98):We sorted the order of the original questions in reverse order.
- run-U3(104):We sorted by score according to the number of page view and click-through rate. (The best run of our group.)
- run-U4(116):We ranked high in the title of the question that contained many white words. We manually selected white words from the title of top 50 questions. The number of target queries was 101 among 1,000 queries in total.
- run-U5(140):We ranked low in the title of the question that contained many black words. We manually selected black words from the title of top 13 questions. The number of target queries was 101.

[3-5] Submitted Runs Cnt'd

Table 2. White words of Run-U4 and black words of RUN-U5.

query ID	query	white words	black words
OLQ-1001	博多	博多駅, ラーメン	博多南線, 彦根城, フラワーアレンジ
OLQ-1002	mt車	AT, クラッチ, 教習所	ジムニーシエラランドベンチャー, 国の検査
OLQ-1003	逃走中	ハンター, 面白い, 台本	房総のむら, 無効票
…	…	…	…

[4] Evaluation

- Table 3 shows submitted run (run name and run ID), Q-measure of offline evaluation, credits of online at the first phase and the second phase.
- The credits are rounded in integers. The '---' stands for no evaluation in the second phase.
- We confirmed the effect of the Page view order as well as OpenLiveQ-1 from run-U2, run9 and run20.
- We also confirmed that the combination of Clickthrough rate that were not very effective with OpenLiveQ-1 became better from run-U3.
- The white and black words methods alone could not achieve the expected effect.
- White words were more effective than black words from run-S2, run-S3, run-N3, run-N4, run-U4 and run-U5.
- The effect could be confirmed by trying both words from run-N6.
- It seems that the combination with page views etc. as run-N7 will increase the effect, but this time we were not able to adjust the combination of parameters.

[4-1] Evaluation Cnt'd

Run		Offline evaluation	Online evaluation (credits)	
Name	ID	Q-measure	First phase	Second phase
run-S0	90	0.38194	-1421	----
run-S1	108	0.42334	-411	---
run-S2	121	0.42256	-428	---
run-S3	132	0.39083	-1253	---
run-S4	151	0.39083	-1254	---
run9	138	0.49021	496	-737
run20	135	0.43063	1039	-211
run-N3	114	0.42346	-1280	---
run-N4	119	0.39556	-961	---
run-N5	142	0.39342	-594	---
run-N6	146	0.41897	-586	---
run-N7	153	0.44076	-310	---
run-U0	94	0.38316	-709	---
run-U1	96	0.49425	424	---
run-U2	98	0.43121	-1411	---
run-U3	104	0.47441	1592	648
run-U4	116	0.38686	-1344	---
run-U5	140	0.38214	-1391	---

Table 3. Runs and their evaluation.

[5] CONCLUSIONS

- Our group OKSAT submitted 18 runs for the NTCIR-14 OpenLiveQ-2 task.
- In this task we reorder questions by using white and black words because most questions in OpenliveQ2-question-data of this task are fit for the queries.
- The white words are selected by the frequency in questions, Google suggest and/or manual.
- On the other hand, the black words are selected by the rareness in the questions.
- The reorder of questions by white words is more effective than by black words from the evaluation results.