HOMEWORK 4: EVALUATION

KEY - OVERVIEW

▸ Calculate annotation consistency

▸ Evaluate performances of Search Engines and your ranking models (VSM, BM25) with MAP and NDCG

▸ Analyze the results in your report

RESULT REQUIREMENTS

▸ For each query (5 queries):

- Kappa - 5 levels

- Kappa - 2 levels

- MAP@10

- NDCG@5

- NDCG@10

- Compress the result files in one ZIP archive

+ Source code (C++, C#, Java, JavaScript, Perl, Python, PHP, Scala, ……)

+ Report (Annotation consistency, SE performance, Your Ranking models performance, Evaluation difficulties,……)

DETAILS

DETAILS

SE	query	description	rank	id	annotation1	anotation	vsm_rank	vsm_score	bm25_rank	bm25_score
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	1	11	2	2				
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	2	12	1					
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	3	13	3	3				
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	4	14	-1					
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	5	15	1	-				
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	6	-1	1	-				
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	7	17	-2					
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	8	18	-2	-2				
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	9	-1	1					
Google	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	10	20	-2					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	1	1	1	2				
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	2	2	1					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	3	3	3	3				
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	4	4	1					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	5	5	3	3				
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	6	6	-1	-				
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	7	7	1					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	8	8	1					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	9	9	1					
Bing	beijing distance to inner mongolia	I want to travel to inner mongolia in the near future for sight seeing. I want to know the distance in order to schedule my visit	10	10	1					

use your results
in Assignment3

rank in SE   Two groups of Annotations

query1/rank.csv

query2/rank.csv

…

query5/rank.csv

KAPPA

Cohen's Kappa coefficient measure for inter-judge (dis)agreement

▸ Each query was annotated 2 times. Kappa = Kappa(A, B)

▸ You need to compute 2 Kappa scores:

- 5-levels Kappa - take annotation result as it is; -2,-1,1,2,3

- 2-levels Kappa - assume -2,-1 as irrelevant; 1,2,3 as relevant

In Assignment2

ANNOTATION DETAILS:

▸ Perfect > Good > Fair > Not relevant > Spam/Page not found

3   >   2   >   1   >   -1   >   -2

MAP

MAP(11-point): Mean 11-point Average Precision

▸ Relevant document: the average of relevance annotation score > 0

▸ Irrelevant document: the average of relevance annotation score <= 0

▸ Set of relevant documents is union of relevant documents from two SE results.

Please note on which side the closed interval is taken

NDCG

Normalized Discounted Cumulative Gain

▸ Document relevant score mapping:

  - R_m is the average of relevance annotation score

  - $R_m >= 2$ map to 3

  - $0 < R_m < 2$ map to 2

  - $-1 < R_m <= 0$ map to 1

  - otherwise map to 0

Please note on which side the closed interval is taken

▸ Compute NDCG@5, NDCG@10

SUBMISSION

▶ File name convention: example

query1/evaluation.csv, query1/rank.csv

query2/evaluation.csv, query1/rank.csv

…

query5/evaluation.csv, query1/rank.csv

    ▶ The file should contain following fields:

- kappa2 - consistency score for 2-levels kappa

- kappa5 - consistency score for 5-levels kappa

- MAP - evaluation score for MAP@10 (2 SEs, 2 models)

- NDCG@5 - evaluation score for NDCG@5 (2 SEs, 2 models)

- NDCG@10 - evaluation score for NDCG@10 (2 SEs, 2 models)

+ Source code (C++, C#, Java, JavaScript, Perl, Python, PHP, Scala, ……)

+report.pdf/docx/…

measure	▶	score
kappa2		
kappa5		
MAP-baidu		
MAP-bing		
MAP-VSM		
MAP-BM25		
NDCG@5-baidu		
NDCG@5-bing		
NDCG@5-VSM		
NDCG@5-BM25		
NDCG@10-baidu		
NDCG@10-bing		
NDCG@10-VSM		
NDCG@10-BM25		

SUBMISSION: REPORT

Introduce your implementation briefly, e.g., how to deal with docID=-1.

Do analysis on:

➢ Annotation consistency

➢ SE performance

➢ Your Ranking model performance

➢ Evaluation difficulties

➢ Other interesting findings and observations

➢ *Please submit the rank.csv with the ranking model score filled in. The results need to be consistent with Assignment3,*

➢ *No requirement on report length. Please include your observations and insights.*

DEADLINE:

23:59 (UTC+08:00, Beijing Time) Apr. 3, Monday
Submit your homework to web learning platform of thu:

Http://learn.tsinghua.edu.cn   our course section "Assignment"(课程作业)

QUESTIONS ?