Groupwise Query Performance Prediction with BERT

Xiaoyang Chen¹,², Ben He¹,², Le Sun²

¹University of Chinese Academy of Sciences; ²Chinese Information Processing Laboratory, Institute of Software, Chinese Academy of Sciences

Experiments

Method

• Encoding Query-Document Pairs:
• Groupwise Prediction:
  • A four-layer transformer enables the cross attention among the [CLS] vectors in each batch.
• Three Different Ranking Context
  • Random Order: all query-document pairs are shuffled.
  • Query Order: position ids are assigned by the initial query order derived by n(σX%).
  • Doc Order: position ids are assigned by initial document ranking.

Overall effectiveness:

• In general, the model incorporating three ranking contexts, i.e., (R+Q+D) outperforms the BERT baselines with all three different model sizes.
• Different ranking contexts leads to different observations on three datasets.

Impact of factors:

• Impact of training batch size: the method works best with a group size of 8.
• Impact of documents per query used in inference: inference with less than 100 documents per query on all three collections yields the best results.

Background

• Query performance prediction (QPP) aims to automatically estimate the search results quality of a given query.
• BERT-QPP[1]: Recent results demonstrate that BERT effectively improves the performance of post-retrieval QPP.
• The groupwise methods have achieved superior performance on learning-to-rank [2] and BERT re-ranking benchmarks [3].
• This paper proposes a BERT-based groupwise QPP method, which simultaneously incorporates the cross-query and cross-document information within an end-to-end learning framework.

References: