中国人民大学 RENMIN UNIVERSITY OF CHINA 高瓴人工智能学院 Gaoling School of Artificial Intelligence

RUCIR at the NTCIR-16 Session Search (SS) Task

Haonan Chen and Zhicheng Dou

Haonan Chen hnchen@ruc.edu.cn

Gaoling School of Artificial Intelligence, Renmin University of China

Background

- To complete a complex search task, a user often has to issue multiple queries to find the information they need.

- Studies have shown that utilizing the contextual information of the search session can help understand the user's current search intent.

Backbone -- COCA

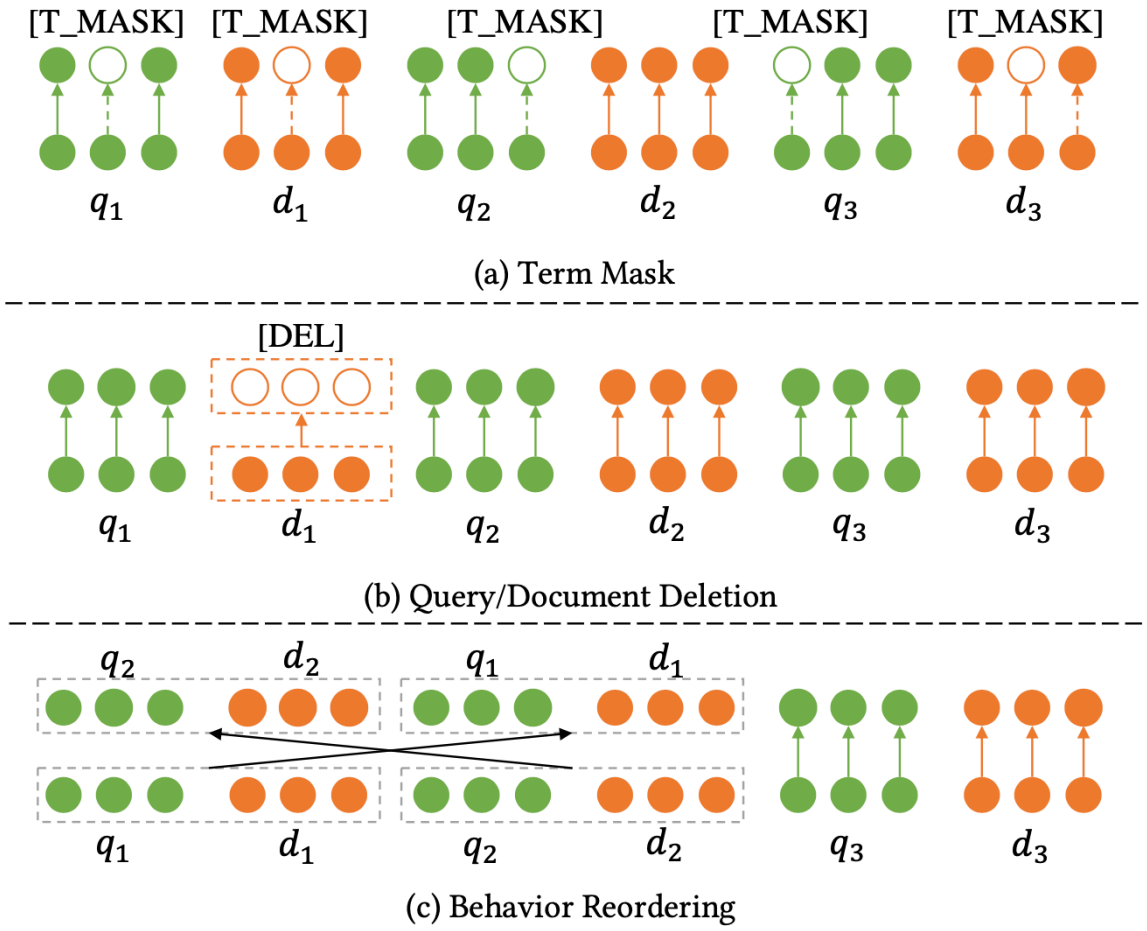
- Using data augmentation strategies to generate possible variations of user behavior sequences from a search log.

- Based on the augmented data, COCA uses contrastive learning to optimize the BERT encoder to train a more robust model to deal with new behavior sequences.

Contrastive Learning of User Behavior Sequence for Context-Aware Document Ranking, Yutao Zhu et al., CIKM 2021.

Data Augmentation Strategies

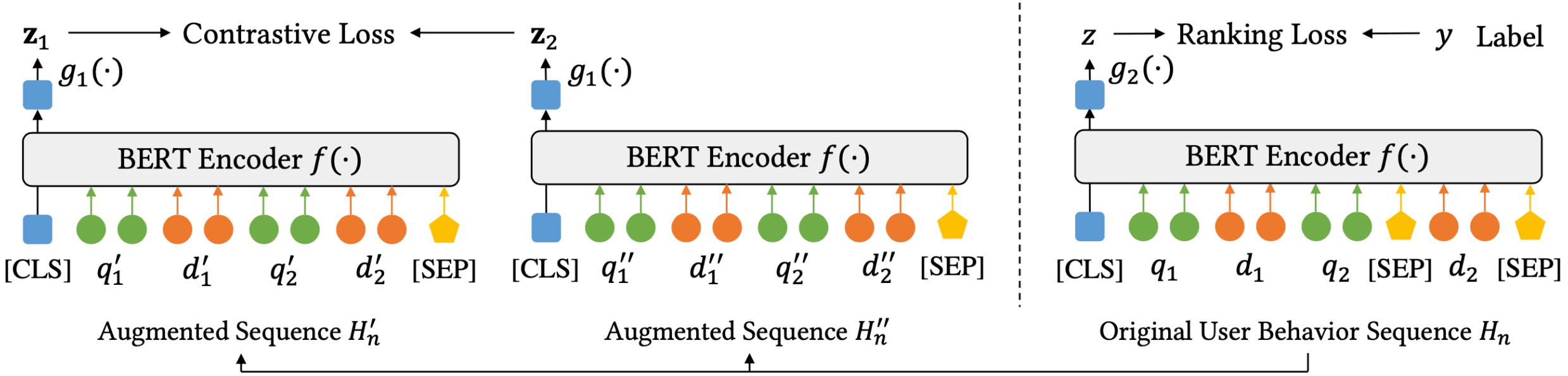
中国人民大学 RENMIN UNIVERSITY OF CHINA 高瓴人工智能学院 Gaoling School of Artificial Intelligence



- Term Mask
- Query/Document Deletion
- Behavior Reordering

Model Architecture

中国人民大学 RENMIN UNIVERSITY OF CHINA 高瓴人工智能学院 Gaoling School of Artificial Intelligence



$$l(i,j) = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{k \neq i} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)},$$

InfoNCE Loss

$$\mathcal{L}_{\text{rank}} = -\frac{1}{N} \sum_{i=1}^{N} y_i \log z_i + (1 - y_i) \log(1 - z_i),$$

CrossEntropy Loss

Some Tricks

- Using BM25 algorithm as regularization.

$$Score(candidate) = \alpha * Score_{COCA} + (1 - \alpha) * Score_{BM25}$$

- Keeping the document that has the largest usefulness value per historical query.

Results

FOSS	**NDCG@3**	**NDCG@5**	**NDCG@10**
COCA+BM25	0.4783	0.4785	0.4939
COCA+U	0.5365	0.5406	0.5570
COCA+BM25+U	**0.5525**	**0.5623**	**0.5693**
POSS	**RsDCG**	**RsRBP**	
COCA	0.4355	0.5640	
COCA+BM25	0.4738	0.6281	
COCA+BM25+U	**0.5439**	**0.7466**	

Thank You for Listening!