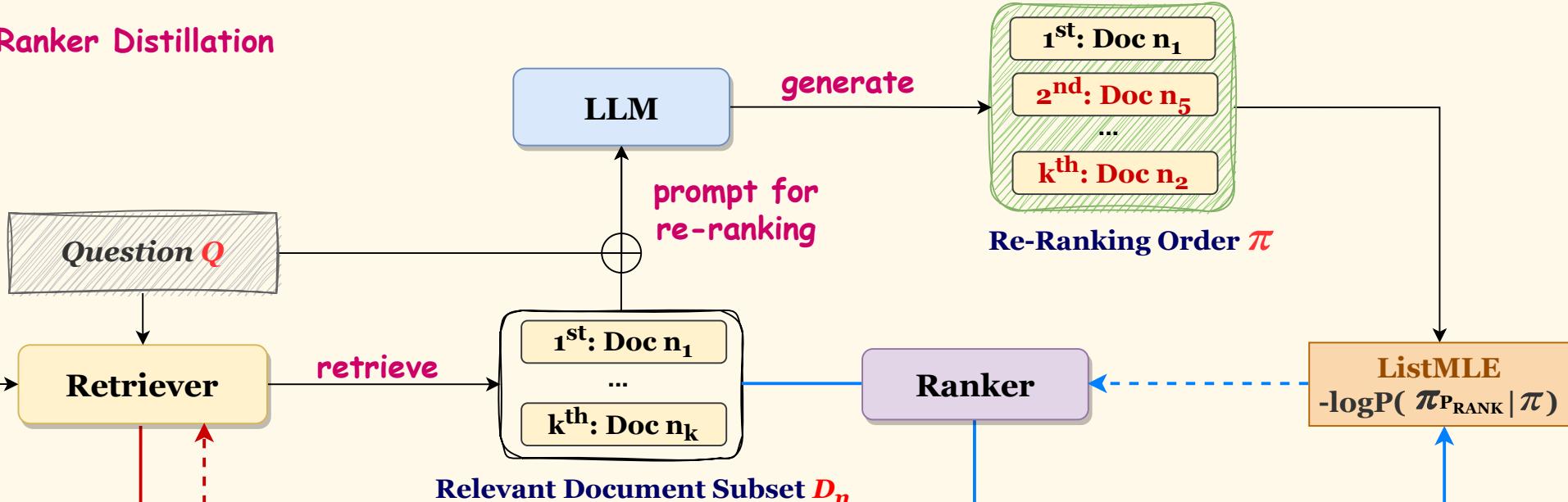


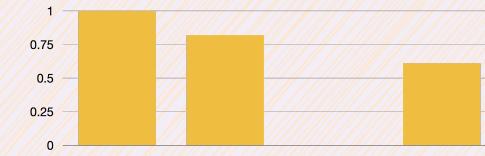
Stage 1: LLM to Ranker Distillation

Knowledge Corpus D

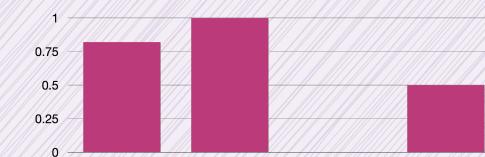


extract

Retriever Similarity Likelihood P_{RETR} over D_n



Ranker Similarity Likelihood P_{RANK} over D_n



extract

KL Divergence
 $KL(P_{\text{RANK}} || P_{\text{RETR}})$

after well-trained

Stage 2: Ranker to Retriever Distillation

- - - distillation training
- data flow of P_{RANK}
- data flow of P_{RETR}