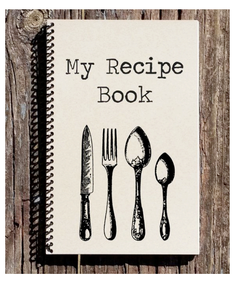
To Interpolate or Not to Interpolate: PRF, Dense and Sparse Retrievers

Hang Li*[1], Shuai Wang*[1], Shengyao Zhuang[1], Ahmed Mourad[1],
Xueguang Ma[2], Jimmy Lin[2], Guido Zuccon[1]
[1] The University of Queensland, [2] University of Waterloo



What's the right recipe?

- Sparse and dense retrievers capture different characteristics of relevance.
- Interpolation can lead to higher retrieval effectiveness.
- Applying interpolation with PRF is unexplored.

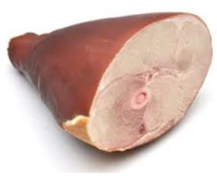
Dense Retriever

- ANCE
- TCT ColBERT V2 HN+
- DistillBERT KD TASB

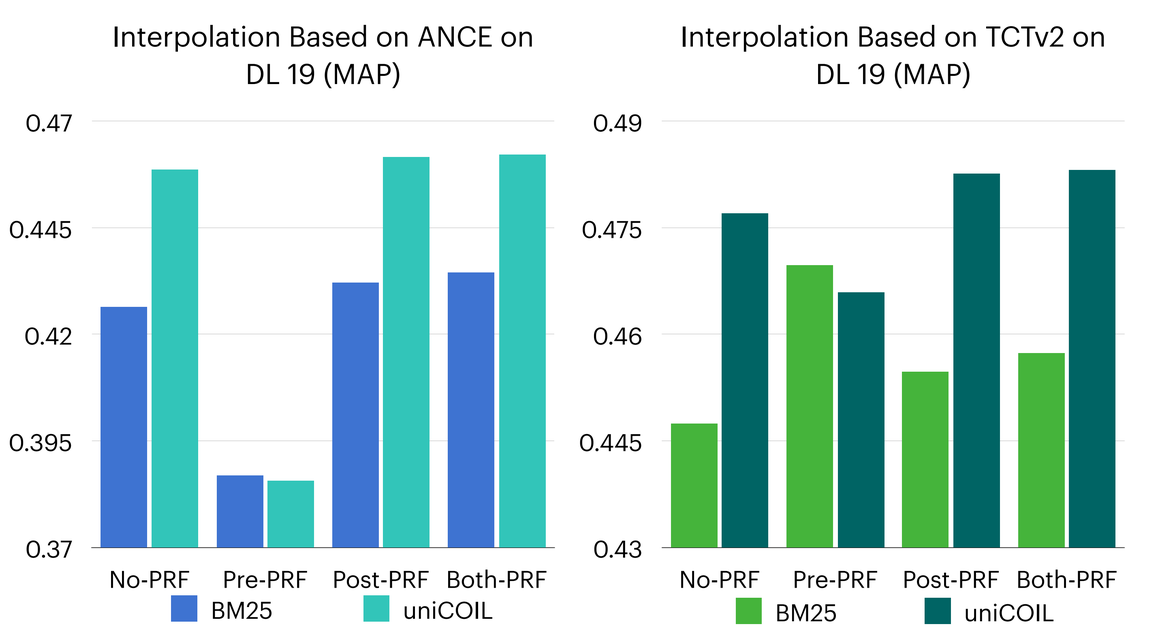
Pseudo Relevance Feedback

- Vector-PRF for dense retrievers

Hawaiian Pizza

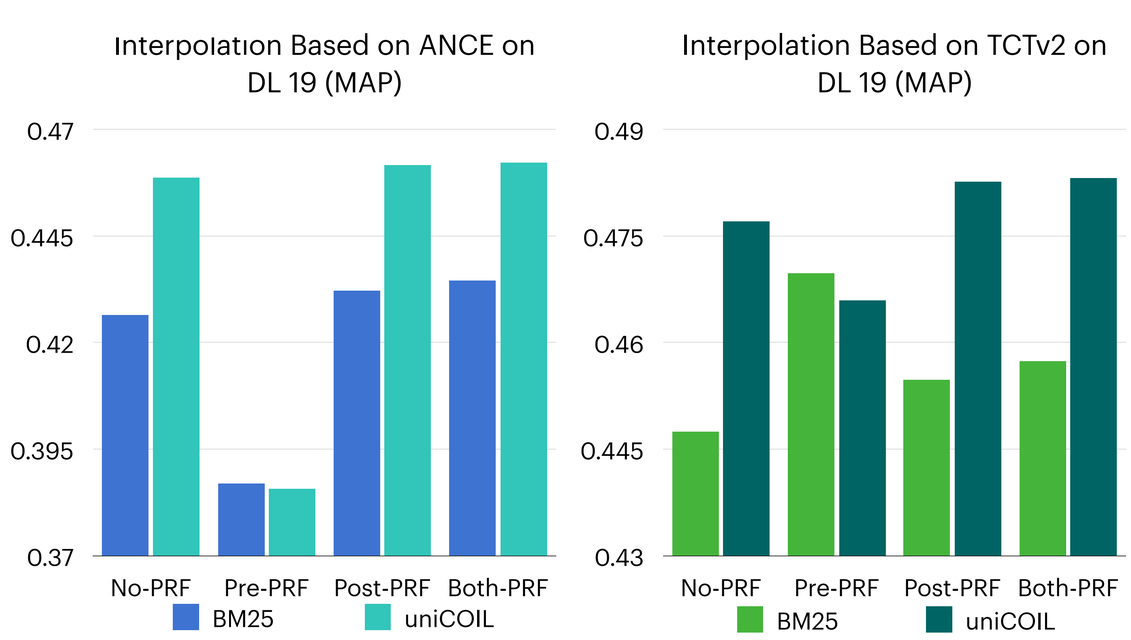


Sparse Retriever: which one?



- Post-PRF & Both-PRF: unsupervised BOWs sparse retriever leads to high recall; neural sparse retriever leads to higher MAP and nDCG@10.
- Pre-PRF: no stable trends wrt sparse retriever.

Interpolation: before or after PRF?



- Interpolation Both-PRF works best.
- Post-PRF has high effectiveness only for uniCOIL sparse retriever.

Find Me

Hang Li is funded by the Grain Research and Development Corporation (GRDC), project AgAsk (UOQ2003-009RTX). Shuai Wang is supported by a UQ Earmarked PhD Scholarship and this research is funded by the Australian Research Council Discovery Projects program ARC DP210104043.

ie lab