Representation performance (ndcg@10)

Model	ndcg@10
llama3-8b	0.479
llama3.1-8b	0.477
llama3-8b-instruct	0.468
llama3.1-8b-instruct	0.464
gte-qwen2-7b-instruct	0.438
qwen2.5-3b-instruct	0.429
qwen2.5-3b	0.427
qwen2.5-7b-instruct	0.427
multilingual-e5-large	0.426
e5-mistral-7b-instruct	0.425
qwen3-4b	0.423
sfr-embedding-mistral	0.422
qwen3-4b-instruct	0.416
deepseek-llm-7b-base	0.409
qwen3-embedding-8b	0.409
multilingual-e5-base	0.406
deepseek-coder-6.7b-instruct	0.399
facebook-contriever	0.397
qwen3-embedding-4b	0.391
deepseek-llm-7b-chat	0.389
qwen3-embedding-0.6b	0.368
e5-base-v2	0.355
e5-large-v2	0.348
bert-base-uncased	0.341
gte-large-en-v1.5	0.339
all-roberta-large-v1	0.329
bge-base-zh-v1.5	0.327
e5-small-v2	0.322
nomic-embed-text-v1.5	0.319
facebook-contriever-msmarco	0.319
bge-large-zh-v1.5	0.317
gte-large	0.313
snowflake-arctic-embed-l-v2	0.312
bge-large-en-v1.5	0.304
multi-qa-mpnet-base-dot-v1	0.304
bge-base-en-v1.5	0.302
mxbai-embed-large-v1	0.297
all-mpnet-base-v2	0.296
distiluse-base-multilingual-cased-v2	0.291
bge-small-en-v1.5	0.288
bge-m3	0.282
all-minilm-l12-v2	0.279
msmarco-distilbert-base-v4	0.275
all-minilm-l6-v2	0.273
allenai-specter	0.269
jina-embeddings-v2-small-en	0.265
paraphrase-multilingual-mpnet-base-v2	0.248
jina-embeddings-v2-base-en	0.097

- - - tfidf

ndcg@10