Representation performance (ndcg@10)

Model	ndcg@10
llama3-8b	0.530
llama3.1-8b	0.529
llama3-8b-instruct	0.517
llama3.1-8b-instruct	0.512
gte-qwen2-7b-instruct	0.492
qwen2.5-7b-instruct	0.479
qwen2.5-3b-instruct	0.475
qwen2.5-3b	0.473
e5-mistral-7b-instruct	0.472
multilingual-e5-large	0.468
sfr-embedding-mistral	0.467
qwen3-4b-instruct	0.466
qwen3-4b	0.465
qwen3-embedding-8b	0.449
facebook-contriever	0.447
qwen3-embedding-4b	0.439
multilingual-e5-base	0.436
qwen3-embedding-0.6b	0.407
e5-base-v2	0.403
bert-base-uncased	0.394
e5-large-v2	0.392
bge-base-zh-v1.5	0.381
gte-large-en-v1.5	0.371
all-roberta-large-v1	0.370
nomic-embed-text-v1	0.366
nomic-embed-text-v1.5	0.364
facebook-contriever-msmarco	0.363
e5-small-v2	0.362
bge-large-zh-v1.5	0.362
gte-large	0.356
bge-base-en-v1.5	0.351
bge-large-en-v1.5	0.338
gte-base	0.334
paraphrase-mpnet-base-v2	0.330
mxbai-embed-large-v1	0.330
all-mpnet-base-v2	0.322
bge-small-en-v1.5	0.321
msmarco-distilbert-base-v4	0.318
all-minilm-l12-v2	0.316
snowflake-arctic-embed-l-v2	0.312
distiluse-base-multilingual-cased-v2	0.306
jina-embeddings-v2-small-en	0.304
allenai-specter	0.292
all-minilm-l6-v2	0.291
bge-m3	0.285
paraphrase-multilingual-mpnet-base-v2	0.253
jina-embeddings-v2-base-en	0.139

tfidf