Work Progress

kNN Search with Parallel Incremental Query Answering

Jaouhara Chanchaf

Tuesday Nov. 22nd, 2022

1. Summary

Done:

AI 1	Prepare for the TOEFL Test.
AI 2	Plot distance, time, precision and recall for query vector NNs.

Not started:

AI 3	Kashif Parallel IQA: cache after each experiment and compare performance for different queries and different threads.
AI 4	In information theory look for some techniques to know when we have enough information (enough NNs).

2. Performance

Experiment over 1GB of data: 100k tables, 494k columns, 5M vectors. 10 Queries of size [50 - 100].

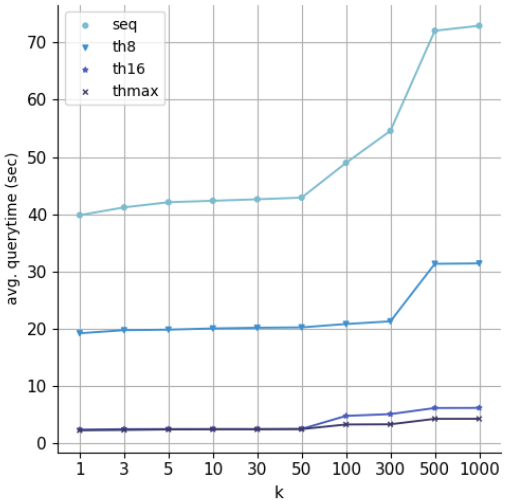

Figure: Kashif Recall

2. Performance

Experiment over 1GB of data: 100k tables, 494k columns, 5M vectors.
1 query of size 91.

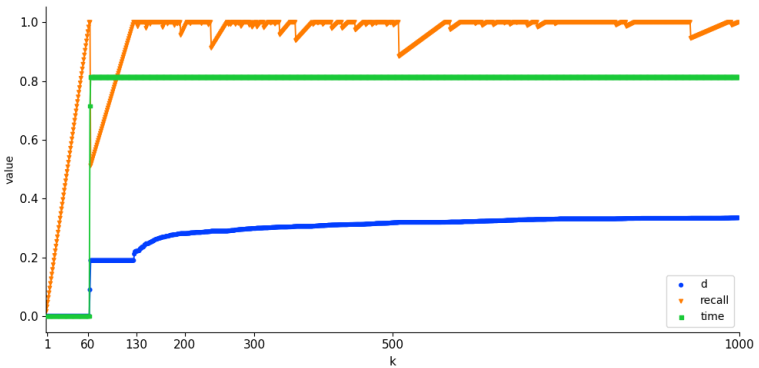

Figure: Kashif: query vector nearest neighbor distance, recall and time.

2. Performance

Experiment over 1GB of data: 100k tables, 494k columns, 5M vectors.
1 query of size 91.

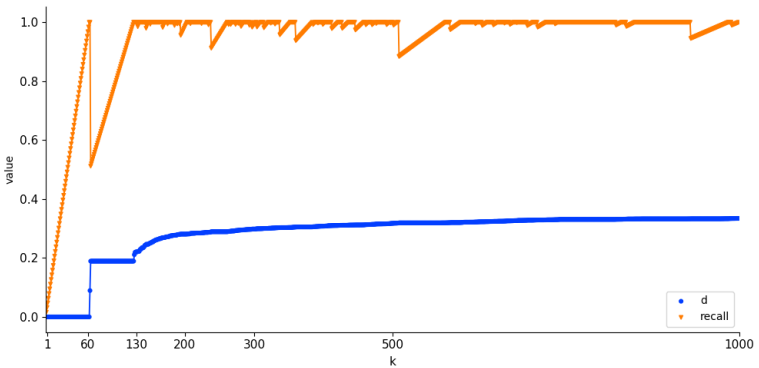

Figure: Kashif: query vector nearest neighbor distance and recall.