Work Progress
kNN Search with Parallel Incremental Query Answering

Jaouhara Chanchaf

Monday Feb 20th, 2023

1. Summary

Done:

AI 2	Measure Kashif recall (using PEXESO as a ground truth) and include it in the query time plots.

In progress:

AI 3	Clean PEXESO and Kashif code and update the Readme files.

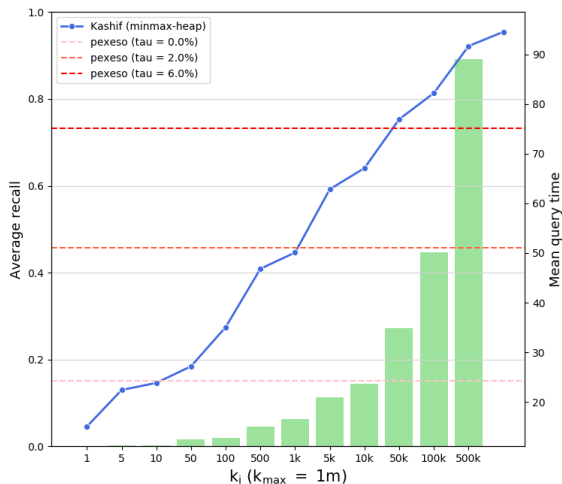# 2. Average recall and mean query time at k_i Kashif (min max heap) Vs PEXESO

Figure 1: Kashif average recall and mean query time
(10 queries, query size = 100, dataset = 100k tables, 490k cols, 5M vectors)

3. Discussion

- Couldn't measure recall at $k_i = 1m$ due to high memory usage.
- Kashif out performs PEXESO ($\tau = 6\%$) in query time for $k_i \leq 50k$.
- For $k_i \leq 50k$ Kashif's recall $\in [0 - 0.27]$.