Work Progress

Demo: KNN Search with Incremental Query Answering

Jaouhara Chanchaf

September 20th, 2022

Summary

Done:

AI 1	Rerun Kashif experiments and measure recall based on vector distance to the query.
AI 2	Add more details to the use case scenarios, result of the second use can should be joinable with the first result from the first use case.
AI 3	Remove query table from list of result.
AI 5	Search for a good use case from the world bank data.

Table: Action Items

In progress:

AI 4	In the demo paper, highlight how higher k values can improve the ranking of relevant tables.

Table: Action Items

Issues:

- To achieve recall ≈ 1, we must choose a very high k value for larger datasets.

Recall and query-time evaluation

Experiment on 100k tables, 494k columns, 5M vectors.
10 Queries of size [50 - 100].

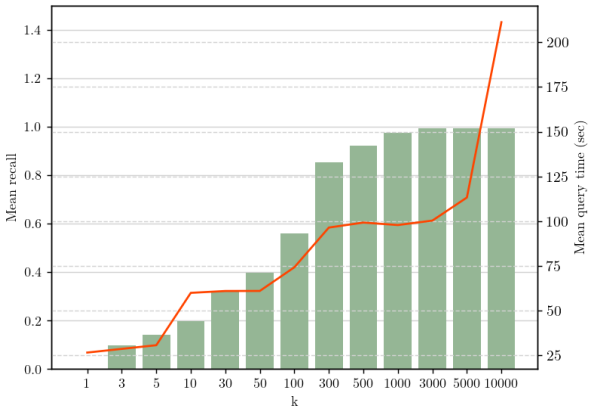

Figure: Kashif Average Recall

Recall

Experiment on 500k tables, 2.47M columns, 25M vectors.
10 Queries of size [50 - 100].

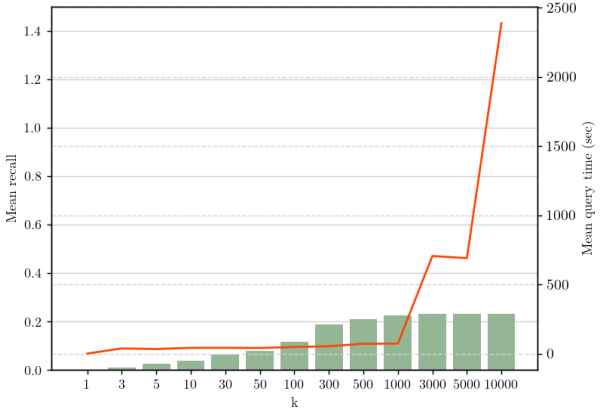

Figure: Kashif Average Recall

Use Case

Keyword Query Result:

Company	Plant	Location	Feedstock	Capacity (MW)
Wheelabrator Technologies Inc.	Wheelabrator Shasta Energy Co. Inc.	Anderson - CA	Logging and Mill Residue/Ag Residue	50
Greenleaf Power LLC	Desert View	Mecca - CA	Ag Residue/Urban Wood Waste	47
Greenleaf Power LLC	Honey Lake	Wendel - CA	Mill and Logging Residue/Forest Thinning/Urban Woodwaste	30
Covanta	Covanta Delano	Delano - CA	Orchard and Vineyard Prunings/Nut Shells/Stone Fruit Pits	58
...	...	...	...	...

Table 1: U.S. Biomass Power Plants

Use Case

Join Query Result:

Category	Plant ID	Plant Name	Unit	Status	Start Date	Retire Date	Prime mover ID	Prime Mover Description	Capacity	net MWh
E	E0027	Desert View Power (Mecca Plant)	GEN1	OP	1991/11/1	-	ST	Steam Turbine	54.15	351291
E	E0041	HL Power Company (Honey Lake)	GEN 1	OP	1989/7/26	-	ST	Steam Turbine	35.5	200712
E	E0029	Covanta Delano, Inc	Delano 1-2	OP	1990/6/12	-	ST	Steam Turbine	58	322731
E	E0086	Wheelabrator Shasta	Units 1-3	OP	1987/1/1	-	ST	Steam Turbine	54.9	405628
...	...	...	...	...	...	...	...	...	...	...

Table 2: Annual Generation - Plant Unit