AWARE-RAG

Group 1

Ness Mayker Chen, Yangxiao Luo, Merve Keskin

Motivation

Problem:

- Generative AI systems produce responses by utilizing the most pertinent data at their disposal.
- Where does this relevant data come from? How applicable is it to the specific context?

Goal:

- Developing a rapid delivery system utilizing external contextual reference data.
- Capable of efficiently searching through and ranking over millions of text in response to user queries.
- Having mechanisms for assessing the qualitative performance of the retrieval outputs.
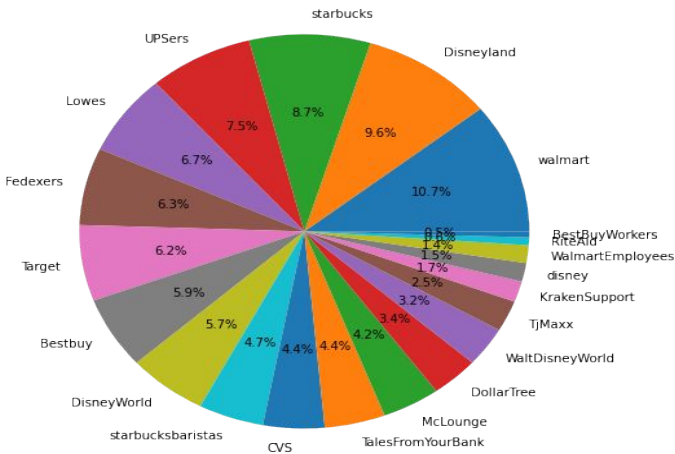

Figure: Hallucinating LLMs, produced by ChatGPT 4

Data

Aware dataset:

- 5 million public Reddit submissions and comments, with 21 subreddits

Challenge: accurately match the comments with submissions

Created a submission-based dataframe with 20613 observations
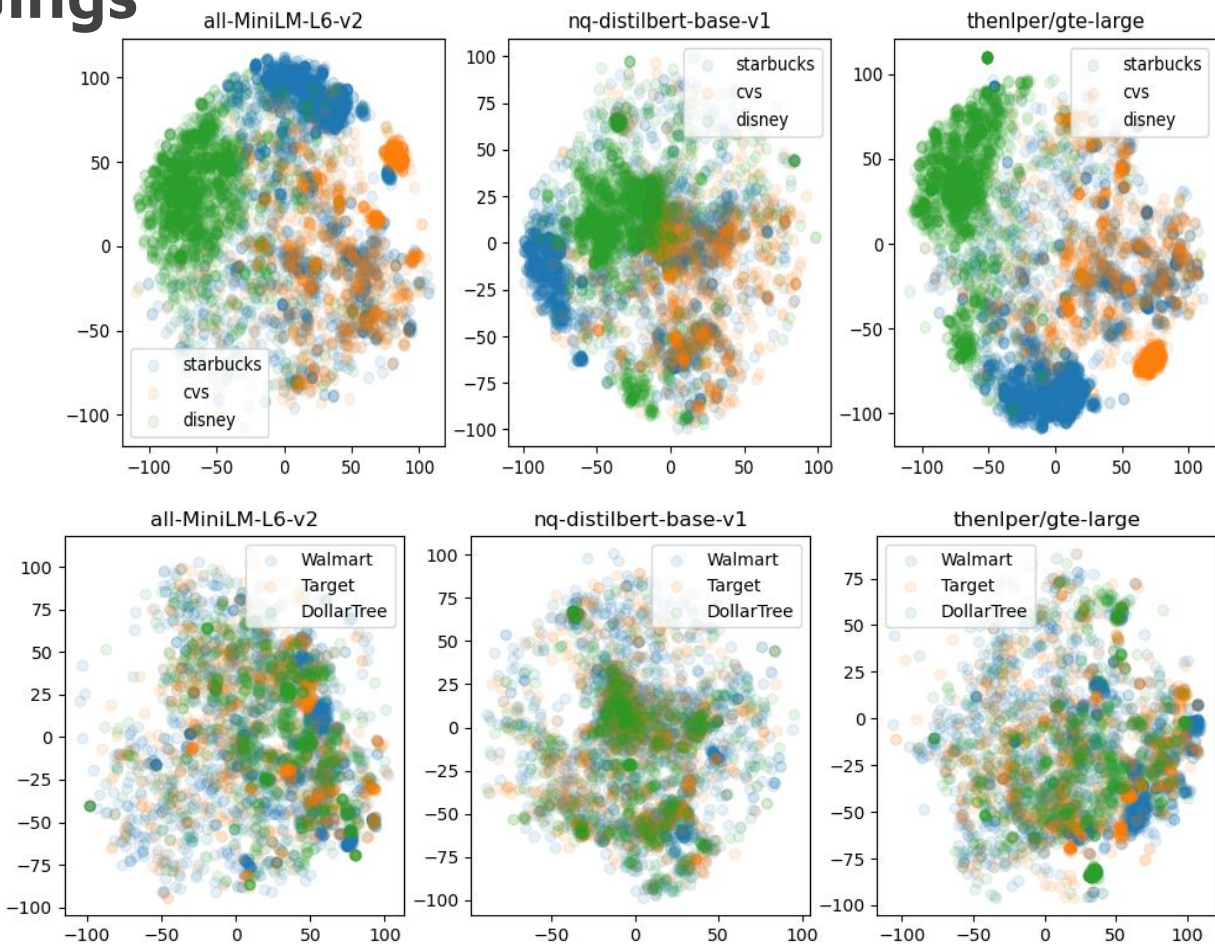
Distribution of Reddit Subreddits

	aware_post_type	aware_created_ts	reddit_id	reddit_name	reddit_created_utc	reddit_author	reddit_text	reddit_permalink	reddit_title	reddit_url	reddit_subreddit	reddit_subreddit_label	index	text_title	text_title_subreddit	comment_indices
0	submission	2023-04-02T13:58:03	129sqka	t3_129sqka	1680458283	MoodyStarGirl	That's it.	/r/starbucks/comments/129sqka/hot_chai_lattes_...	Hot chai lattes shouldn't have water	https://www.reddit.com/r/starbucks/comments/12...	starbucks	0	0	Hot chai lattes shouldn't have water [SEP] Tha...	Hot chai lattes shouldn't have water That's it...	[1, 5, 12, 13]
1	submission	2023-04-02T14:07:43	129t0vd	t3_129t0vd	1680458863	kotakehoe	i had the urge to make a cute baby frapp in a ...	/r/starbucks/comments/129t0vd/my_cute_little_f...	my cute little frappucino :)	https://i.redd.it/0gd8tk2zyjra1.jpg	starbucks	0	18	my cute little frappucino :) [SEP] i had the u...	my cute little frappucino :) i had the urge to...	[19, 20]
2	submission	2023-04-02T14:20:39	129te77	t3_129te77	1680459639	Extension_Cricket_30	Hey guys! I was wondering what benefits people...	/r/starbucks/comments/129te77/partner_benefits...	Partner Benefits at your Starbucks?	https://www.reddit.com/r/starbucks/comments/12...	starbucks	0	26	Partner Benefits at your Starbucks? [SEP] Hey ...	Partner Benefits at your Starbucks? Hey guys! ...	[27, 28, 29]
3	submission	2023-04-02T14:39:11	129txfg	t3_129txfg	1680460751	ateez-lvr	did anyone notice (or is it the same for you a...	/r/starbucks/comments/129txfg/not_really_a_ran...	not really a rant, just an observation	https://www.reddit.com/r/starbucks/comments/12...	starbucks	0	30	not really a rant, just an observation [SEP] d...	not really a rant, just an observation did any...	[31, 37]
4	submission	2023-04-02T14:43:36	129u231	t3_129u231	1680461016	chuchae	hi friends, so i recently got back from a we...	/r/starbucks/comments/129u231/accidental_no_show/	accidental no show	https://www.reddit.com/r/starbucks/comments/12...	starbucks	0	38	accidental no show [SEP] hi friends, so i re...	accidental no show hi friends, so i recently...	[39]
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
23263	submission	2023-04-16T10:57:21	12obutt	t3_12obutt	1681657041	_TheBiggestSad_	My family and I are going to Disney World in a...	/r/WaltDisneyWorld/comments/12obutt/cheap_plac...	Cheap places to eat in/out of Disney	https://www.reddit.com/r/WaltDisneyWorld/comme...	WaltDisneyWorld	20	404184	Cheap places to eat in/out of Disney [SEP] My ...	Cheap places to eat in/out of Disney My family...	[404185, 404187, 404188, 404189, 404189, 404189,...]
23264	submission	2023-04-16T10:03:41	12o9gli	t3_12o9gli	1681653821	purplechunkymonkey	My daughter and I are doing a girls trip. She ...	/r/WaltDisneyWorld/comments/12o9gli/what_to_do...	What to do outside of the parks	https://www.reddit.com/r/WaltDisneyWorld/comme...	WaltDisneyWorld	20	404248	What to do outside of the parks [SEP] My daugh...	What to do outside of the parks My daughter an...	[404249, 404250, 404251, 404252, 404253]
23265	submission	2023-04-16T09:20:24	12o89v3	t3_12o89v3	1681651224	ol_dirty_applesauce	Shopdisney.com only has the sleeveless version...	/r/WaltDisneyWorld/comments/12o89v3/is_this_to...	Is this top offered as a t-shirt or sweatshirt...	https://i.redd.it/1bf4q46hgaua1.jpg	WaltDisneyWorld	20	404258	Is this top offered as a t-shirt or sweatshirt...	Is this top offered as a t-shirt or sweatshirt...	[404259]
23266	submission	2023-04-16T09:18:20	12o87ug	t3_12o87ug	1681651100	Savings_Spell6563	Upon Splash Mountain closing, I would have re-...	/r/WaltDisneyWorld/comments/12o87ug/if_you_had...	If you had the power to go back in time and ch...	https://www.reddit.com/r/WaltDisneyWorld/comme...	WaltDisneyWorld	20	404261	If you had the power to go back in time and ch...	If you had the power to go back in time and ch...	[404262, 404263, 404264, 404265, 404266, 404266,...]
23267	submission	2023-04-16T08:59:13	12o7ozd	t3_12o7ozd	1681649953	bubblekitty607	Hi, so I want to buy the pixie dust annual pa...	/r/WaltDisneyWorld/comments/12o7ozd/all_4_park...	All 4 parks, one per day - possible for pixie ...	https://www.reddit.com/r/WaltDisneyWorld/comme...	WaltDisneyWorld	20	404676	All 4 parks, one per day - possible for pixie ...	All 4 parks, one per day - possible for pixie ...	[404677, 404679, 404680, 404681]

20613 rows × 16 columns
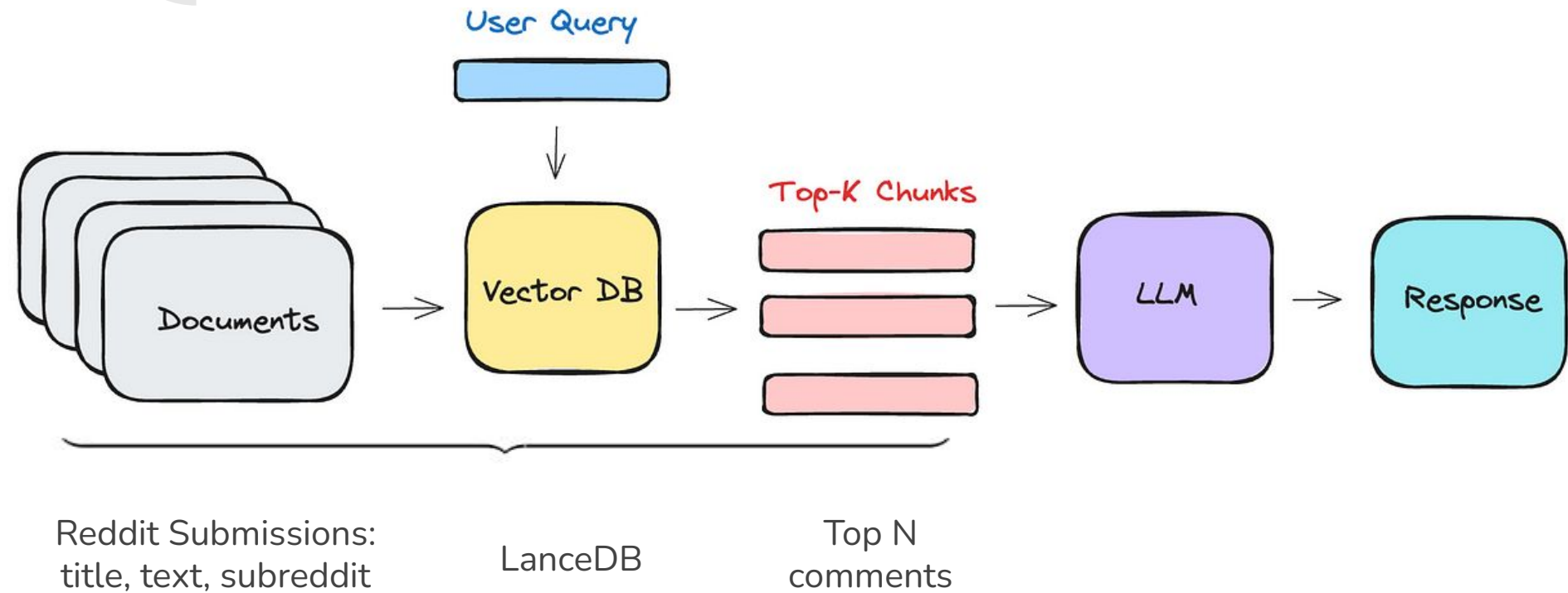
Visualizing Embeddings with t-SNE

We test three embeddings for our RAG system.

All three show similar degrees of 2D separation when subreddits topics are more diverse (top panel) and when subreddit topics are more similar (bottom panel).

We use "all-MiniLM-L6-v2" for our embeddings due to its small size and fast performance.

Outline of Pipeline

User Query

Documents

Vector DB

Top-K Chunks

LLM

Response

Reddit Submissions:
title, text, subreddit

LanceDB

Top N
comments

Retrieval approach

- For each submission s, encode its "reddit_title [SEP] reddit_text # subreddit" to obtain a vector v_s.

- Given a query q, use the same embedding to encode it to obtain a vector v_q.

- Use KNN to find $v_{s_1}, v_{s_2}, ..., v_{s_n}$, the n submissions closest to v_q.

- Return the most upvoted comments (contexts) $c_1, c_2, ..., c_n$ of the submissions $s_1, s_2, ..., s_n$

- We evaluated the model with embedding = "all-miniLM-L6-v2" and $n = 1$.

Evaluation

Answers generated by llama 3-8B with no contexts

	question	ground_truth	answer	answer_relevancy
0	Is my vacation time messed up? Start date Oct ...	Your union steward will be a better resource t...	When you start as a PTF, your Seniority date ...	0.794721
1	Coordinator Position [SEP] I officially submit...	Yeah, expect to do a *lot* more work for not a...	1. The first thing to remember is that this is...	0.756197
2	our store schedules us until 6 to cut hours, b...	You only need to cover what you're scheduled. ...	2) Yes, you don't have to put in more than you...	0.805190
3	Terrible day one experience - I would not advi...	Hi u/grissinb 👋, Val from Kraken here. I see...	1. As a user I can't confirm this happened as ...	0.747089
4	How to change survey. [SEP] Long long ago...(201...	Rite Aid surveys aren't about pleasing the cli...	1. We had a survey for our customers and we go...	0.000000
5	I finally bit the bullet and promoted myself t...	Congratulations on the promotion! Seriously th...	So you're one of them people who wants a job ...	0.000000

	answer_relevancy
count	50.000000
mean	0.560698
std	0.354931
min	0.000000
25%	0.000000
50%	0.749884
75%	0.801933
max	0.842673

Answers generated by llama 3-8B with contexts given by our RAG system

	question	contexts	ground_truth	answer	answer_relevancy
0	Is my vacation time messed up? Start date Oct ...	['You can find your seniority date on the UPSe...	Your union steward will be a better resource t...	Yes, your PT service does count towards FT se...	0.796527
1	Coordinator Position [SEP] I officially submit...	['Which coordinator? there's 4 different types...	Yeah, expect to do a *lot* more work for not a...	1st off, congrats on making it so far in the p...	0.756164
2	our store schedules us until 6 to cut hours, b...	['If you see a . (you know, a *decimal*) it me...	You only need to cover what you're scheduled. ...	It's not a big deal if you do it wrong. You d...	0.000000
3	Terrible day one experience - I would not advi...	["Hello u/Butterscotch766, Safeguarding your ...	Hi u/grissinb 👋, Val from Kraken here. I see...	I'm sorry for your experience on our platform...	0.738978
4	How to change survey. [SEP] Long long ago...(201...	['They don't listen to it anyway lmfao'\n 'The...	Rite Aid surveys aren't about pleasing the cli...	2 questions that are more important than the o...	0.730379
5	I finally bit the bullet and promoted myself t...	['I cannot wait to quit, or to at least leave ...	Congratulations on the promotion! Seriously th...	[SEP] After nearly 2.3 years working at this ...	0.864723

	answer_relevancy
count	50.000000
mean	0.613419
std	0.331259
min	0.000000
25%	0.730622
50%	0.764282
75%	0.799226
max	0.893584

Discussion and Next Steps

Use complete documents and experiment with different chunk sizes

Explore alternative metrics

Define an "I don't know" threshold on the RAG side of the model

Thank you Erdös and Aware!

We want to give a big thank you for this wonderful opportunity to learn and improve our data science skills!