SIGNAL

Search-**I**nfluenced **G**raphical **N**arrative of **A** **L**ife

Objective

- Separate nuanced signals from noise

- Create User profile

RAG (Retrieval-Augmented Generation)

- Clean queries

- Embed queries (Dense Vector)

- For new question about user's profile

    a. Retrieve top n similar questions as context

    b. Answer based on context

RAG (Retrieval-Augmented Generation)

Cons

- Nuanced topics away from general topics, risk being treated as noise

Clustering + Topic Labeling

- Clean queries

- Embed queries (Dense Vector)

- Cluster embeddings

- Label clusters (LLM or keyword-based)

- Filter noise (Remove small clusters)

- Build user profile(Store the label, top queries, timestamps, and cluster embedding (centroid).)

Clustering + Topic Labeling

Cons

- Nuanced topic's embedding far from centroid, treated as noise

Knowledge Graph

- Clean queries

- Extract entities(NER/LLM)

- Extract relationships(LLM)

- Graph: Nodes > Entity; Edge > Relation

- Remove nodes with few connections

- When answering question convert it into graph

- Traverse KG starting from nodes present in question

- Use retrieved questions as context to answer user's question

Knowledge Graph

Cons

- No explicit connection between nodes of nuanced topics to central themes

SIGNAL

- Clean queries

- Retrieve context: use RAG to fetch similar prior searches within a time window

- *Filtration*: LLM extracts central themes from searches(past week) that influence future behavior.

- Build entity graphs: LLM extracts entities and relationships from each query.

- Convert themes and entities into graph nodes and relation edges for a Knowledge Graph (KG)

- *Nuance*: embed node labels and merge semantically similar nodes for generalization.

- Repeat for each query, updating the KG dynamically over regular intervals.

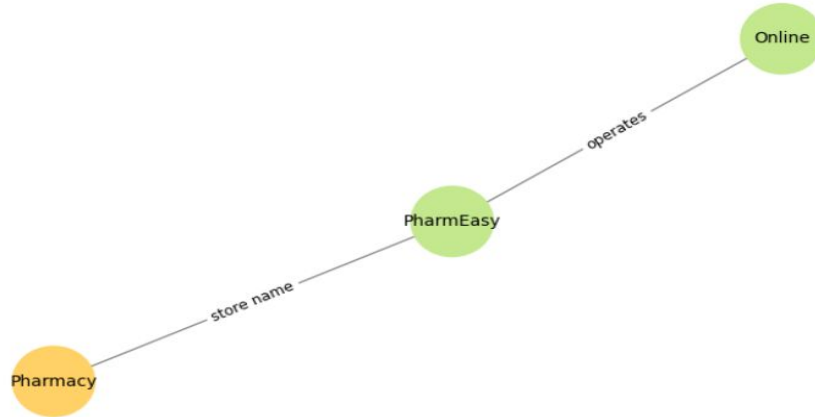- *Cleaning* removes clusters with size less than threshold with no recent addition

Filtration

- For each query retrieve question from _past week_ that are _similar_ to it

- _Past Week:_ Small window ensures (_i._) relevance, (_ii._) reduced computation, (_iii._) yet large enough for LLM to efficiently generalize over evolving topics (**Assumption:** In 1 week specialization within a theme would be gradual enough that the llm will be able to infer the general theme)

- _Similar_: Use sentence embedding to get dense vector and retrieve similar(_Future Influencing_) questions from past week (**Assumption:** In 1 week specialization within a theme would be gradual enough that the similarity search would be able to group them together)

- Only queries passing this _Future Influencing_ (**Assumption**: Non future Influencing questions are considered noise) filter enter the knowledge graph pipeline.

- LLM generalizes the influential queries into high-level themes; non-influential queries are ignored.

- **Dynamic noise reconsideration**: queries initially considered noise can enter the graph if they later influence subsequent searches.

Filtration

Question 5: Searched for pharmeasy trustpilot
Retrieved questions:
- Visited PharmEasy: Online Pharmacy & Medical Store in India | 50 Lakhs+ ...
- Searched for india online pharmacy
- Searched for india best online pharmacy
- Visited Online Medical Store: Order Medicine Online & Get Fastest Delivery ...

- Using only searches that influence the future to create theme nodes allows for only strong signal to be involved in creation of theme nodes while ignoring noise
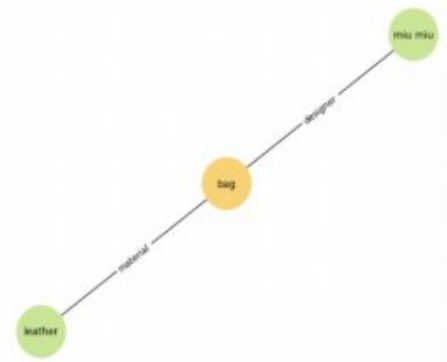
Nuance

- For every new node calculate similarity(sim) with all existing nodes;

- if sim ≥ node_th1 _generalize into 1 label_ with n hop neighbors as context

- if node_th2 < sim < node_th1 _connect via edge_ as related and use searches from n hop neighbours as context

- _Nuance_: preserve nuanced nodes as leaves

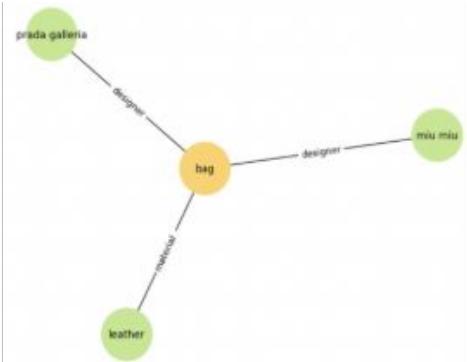- _Generalization_: combine nodes to capture the overarching, nuance interest under a single general theme as root nodes.

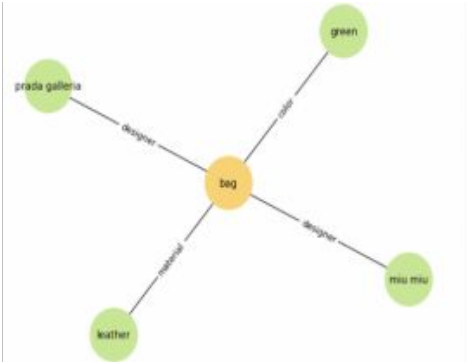_Please Note: The following page shows first 8 specialization of one theme_
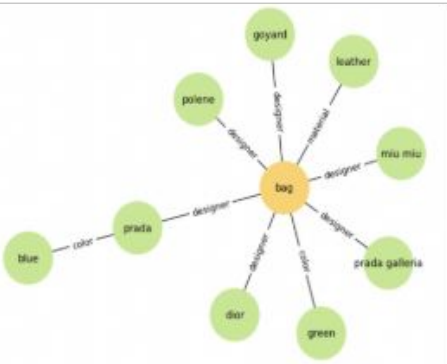
Nuance
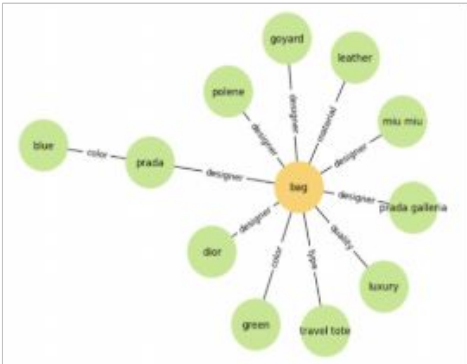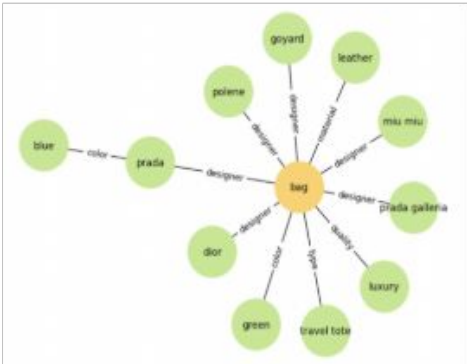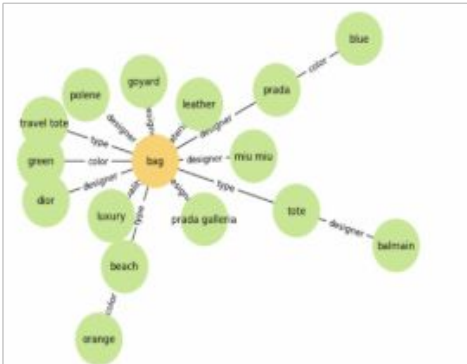
Cleaning

- **Why Necessary**: Filtration allows short term interests to create theme nodes,

- Remove nodes if they have not evolved into long term interests(_cluster size_ less than threshold) over a given _timeframe_

Pros

- 2 explicit methods to connect nuanced nodes to central themes:

  a. *Filtration* builds the central theme using similar questions within a short timeframe to capture gradual specialization.

  b. *Nuance* generalizes central themes (sim>node_th1) connects them if closely related (node_th2<sim<node_th1)

- 2 explicit methods to remove noise

  a. *Filtration* distinguishes future influencing searches from those that don't

  b. *Cleaning* allows *short* term interests to emerge while removing those that don't persist.

Cons

- **Threshold sensitivity**: node merging and similarity thresholds need careful tuning.

- **Windowing issue:** past-week may miss short, intense bursts of searches(deep specialization in short period of time); experiment with fixed-number-of-searches windows instead.

- **Noise Removal:** _Clusters size_ less than threshold when not used for the _past year_ removal, needs careful tuning, since it includes _permanent removal_