ANALYSIS

Part 1: Semantic Search Model

What type of queries tend to do well? Which is not so well?

Effective queries are usually those that are clear, precise, and closely match the content of the database. When terms or phrases are used in search searches, they yield accurate results, for instance, they are expected to be in the film's title or plot description. These inquiries are usually straightforward and contain specifics about the film, such as names of characters, locations, or standout plot aspects.

Conversely, generic, unclear, or having a broad scope of movie genres covered in their wording are examples of less effective queries. These queries aren't specific enough to differentiate between different movies in the library. It would be challenging for the model to decide which romantic movie from the 1920s is the best choice in the absence of further precise information if the query "romantic movie from the 1920s" applied to a variety of different films.

For the queries that the model didn’t perform well, what could be two alternative approaches?

For questions 3 and 4, the model did not perform well, returning four irrelevant and only one relevant search result. It's because of providing every exact detail in the questions. Take the sample query "Love story, good ending takes place in France," for instance. This type of search will provide French films with happy endings and romantic storylines; nevertheless, only relevant movie plots contain all three of these components. Thus, the outcomes—which don't include only France, love tales, or unhappy endings—don't matter. After employing the "all-MiniLM-L6-v2" model technique, the issue appears. Other strategies that we could employ to address the issue include.

User Feedback loop: Implementing a feedback loop that lets users rank the importance of search results is another way to improve performance. By using this input over time to learn from the questions that initially performed poorly, the search algorithm can be made even more effective. By analyzing the common characteristics of searches that are consistently marked as irrelevant, the model can adjust its search parameters to better match user expectations.

Interactive Query Refinement:

Provide an interactive process for refining queries so that the user can provide more details or explanation if the initial query is too general or yields poor results. This tactic can include requesting the user to choose from a list of related topics or suggesting more specific terms to focus the search. In addition to improving the search results, this method assists in locating common patterns in questions that the model finds difficult to handle.

Execution Summary

Recommendation:

Semantic search:

1. Benefits:

- Faster and more straightforward to use.
- Can scale for huge datasets and be easily implemented.

Cons: - Compared to RAG and Re-Ranked combining methods, accuracy may not be significantly higher. The quality of the embeddings is crucial.

- It has trouble processing data when there are several restrictions (specific queries) applied, such as "Thriller, Germany, Dark."

2. Re-Ranked combining: • Benefits: - It combines ranked concepts with the advantages of Semantic Search and BM25.

- When used with Cross Encoders, it can even manage numerous restrictions in the query.

Cons: Managing and integrating several models requires more complexity and computational power. It takes a long time. The efficiency of the result combination relies on the initial retrieval techniques employed**.**

3. Recovery Augmented Generation, or RAG:

The model's primary idea of sending a pertinent document along with a query to improve results is one of its main advantages. It also understands context, which contributes to its high accuracy.

- When compared to traditional generation methods, can yield results that are more contextually relevant and instructive.

Cons: -High processing power and a more intricate architecture are needed. It is very time-consuming.

- It is exceedingly intricate and challenging to put into practice, particularly when using the idea with a manually uploaded dataset as opposed to an online dataset.

We would suggest using the "Re- Ranked Methods" approach because Semantic Search was a more straightforward and accurate model. However, given the complexity of real-world datasets, we think real-world models need to be a little more sophisticated to access and search data efficiently. Implementing RAG was the most challenging of the RAG and Re Ranked methods while working with the manually loaded dataset, but it won't be a major issue because there will be a lot of datasets available online. But there would be a terrible disaster if the server broke down. Semantic search (BM25) is covered by the reranked idea.

Production:

The model works well since it makes obtaining the findings simple. However, there are some extra requirements to implement it in the real world. These include selecting a cloud infrastructure to be deployed, such as AWS, Google Cloud, or Azure, and any model needs to have updates and the concept of reusability implemented. We didn't upload everything in accordance with the guidelines. To enable everyone to understand and attempt to duplicate it, everything needs to be made crystal obvious.

Scalability may be an issue for these tasks because of the size of the datasets; to address this, we want to distribute the load across several servers to achieve load balancing. While managing large datasets may be difficult and costly, resources for monitoring will help save some money. Nevertheless, we advise using the "Re-Ranked" model because it will ultimately turn a profit and is still more economical than RAG.

RAGs/ LLMs:

How can the performance of your model be assessed? How can it be compared to BM25, Semantic Search, and reranked? elements influencing the performance of the models

The model's output gives us insight into how well it performs. It produced pertinent outcomes despite its complexity. With the exception of the second and fifth queries, nearly all results provided three pertinent results. Both questions had a distinct plot and were quite specific. It provided at least one useful result, however for a single case, BM25 produced all irrelevant results.

A model's performance can be assessed using a number of metrics, including F1 Score, Precision@k, Recall@k, and MMR.We solely worked on Recall@K and MMR for this project. When we used the metrics and pertinent search results to examine these models. Compared to BM 25 and Semantic Search, it was superior. Although the RAG theoretical concept was easy, it was challenging to apply using a manual dataset from a local device, which resulted in a throat struggle when it came to re-ranking. The model's effectiveness may be impacted by unclear queries and a deficiency of five relevant search results. Additionally, using suitable techniques like Mistral and directly downloading datasets from the internet would have produced different outcomes.

Fine Tuning:

One can use a variety of pre-trained models in sentence transformer libraries such as BERT, among others. The issue is that every model is trained using a specific dataset to achieve a specific assignment. If we use a single pre-trained model and it hasn't been trained on the dataset previously, we will lose a significant amount of accuracy. Fine Tuning fills that need. Fine-tuning entails training the specific model on the dataset we intend to use so that it becomes familiar with the dataset and, having previously trained on that domain, can provide higher accuracy when we perform the task. This allows us to employ any model with higher accuracy for our project, if we fine-tune it beforehand so that it has some understanding of our domain before being used in the project.

APPENDIX

Part 1: Semantic Search Model

Query 1: “Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions”

1. Nanook of the North
2. Kizhakkunarum Pakshi
3. The 11th Hour
4. O Pioneers!
5. Being Caribou

Query 2: “Western romance”

1. A Saloon Wet With Beautiful Women
2. Lal Dupatta Malmal Ka
3. Goodnight, Sweetheart
4. Book of love
5. Modern Romance

Query 3: "Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."

1. Picture Brides
2. A korean in paris
3. Golden Arrow
4. Desert Legion
5. Hotel Chelsea

Query 4: "Comedy film, office disguises, boss's daughter, elopement."

1. Lal Dupatta Malmal Ka
2. Ami Aar Amar GirlFriends
3. The boy friends
4. Look Up Your Daughters!
5. Interlude

Query 5: "Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."

1. Cleopatra
2. Caesar and Cleopatra
3. Loves Labours lost
4. Lost and Found
5. Royal Treasure

Query 6: "Denis Gage Deane-Tanner"

1. The baxter
2. Meet Simon Cherry
3. Weddings Rehearsal
4. Boys Are Back, the boys are back

5. The running man

Part 2: Reranker notebook

1. **BM25 Method**

Query 1: “Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions”

1. Papilio Buddha
2. Not With my wife, you Don’t
3. Uriyadi
4. Ladies
5. Being Caribou

Query 2: “Western romance”

1. Hyderabad Blues
2. Out of this world
3. El khar
4. Blightly
5. The big show

Query 3: "Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."

1. Sahara
2. Inspiration
3. Ruby Cairo
4. Medusa
5. He who gets slapped

Query 4: "Comedy film, office disguises, boss's daughter, elopement."

1. Floot tide
2. Class Of Nuke ‘Em High 3
3. The Milkman
4. Delete my Love
5. The Boy Friend

Query 5: "Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."

1. Cleopatra
2. Mama’s Affair
3. Peter Pan
4. Madame X
5. This Exquisite Thief

Query 6: "Denis Gage Deane-Tanner"

1. Captain Alvarez
2. Indecent Proposal

3. I love you, Beth Cooper
4. Yolki 6
5. The Stranger door

2. BM-25 and Retrieval, Re-ranked both combined

Query 1: “Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions”

1. The Savage Innocents
2. The Last Wave
3. Nanook of the north
4. Searchers
5. The White Dawn

Query 2: “Western romance”

1. El Akhar
2. The Big Show
3. The Magic of Belle Isle
4. Saturdays millions
5. This is not what I expected

Query 3: "Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."

1. Sahara
2. Death on the nile
3. He Who Gets Slapped
4. Cario time
5. The Suburanite

Query 4: "Comedy film, office disguises, boss's daughter, elopement."

1. Mabels brother
2. Bucking Broadway
3. the double
4. He hired the boss
5.

Query 5: "Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."

1. Cleopatra
2. Mama’s Affair
3. The Hunchback of Notre Dame
4. Three Ages
5. What Daisy Said

Query 6: "Denis Gage Deane-Tanner"

1. Captain Alvarez

2. I Love You, Beth Cooper
3. Skin & bone
4. Dean
5. Indecent proposal

Part 3: Retrieval Augmented Generation (RAG)

Query 1: “Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions”

1. The white Dawn
2. The savage Innocents
3. Nanook of North
4. Born in china
5. The forzen limits

Query 2: “Western romance”

1.Kalgejje
2.A saloon wet with Beautiful Women
3.Age Jodi Jantam Tui Hobi Por
4.Mudhal Kadhal Mazhai
5.Graduate

Query 3: "Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."

1.Sahara
2.Youre my everthing
3.Night terrors
4.Shangai Story
5.Mee Sindhutai Sapkal

Query 4: "Comedy film, office disguises, boss's daughter, elopement."

1.Amarilly of Clothes-line alley
2.Manmagan thevai
3.ask father
4.Sweety nanna jodi
5.He Laughed Last

Query 5: "Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."

1.the wrestler
2.another midnight run
3.last of the comanches
4.captain Alvarez
5.The skywayman

Query 6: "Denis Gage Deane-Tanner"

1. Captain Alvarez
2. Near the Rainbow's
3. A Man from Wyoming
4. The Wolf Song
5. Tenderloin

Part 4: All in one Data Table

Query 1	Recall@1	Mean Reciprocal Rank
Semantic Search	0.33	1
BM25	0	0
Reranker	0.33	1
RAG	0.33	1

Query 2	Recall@1	Mean Reciprocal Rank
Semantic Search	0.33	1
BM25	0.25	1
Reranker	0.33	1
RAG	0.33	1

Query 3	Recall@1	Mean Reciprocal Rank
Semantic Search	1	1
BM25	0.5	1
Reranker	0.5	1
RAG	0.5	1

Query 4	Recall@1	Mean Reciprocal Rank
Semantic Search	0.5	1
BM25	0	2
Reranker	0.5	1
RAG	0.5	1

Query 5	Recall@1	Mean Reciprocal Rank
Semantic Search	0.5	1
BM25	1	1
Reranker	1	1
RAG	1	1

Query 6	Recall@1	Mean Reciprocal Rank
Semantic Search	0.5	1
BM25	1	1
Reranker	0.5	1
RAG	0.5	1